



## INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

<b>(51) International Patent Classification <sup>7</sup> :</b> <b>C12N 15/12, C07K 14/47, A61K 31/70,</b> <b>35/12, 38/13, C07K 16/18, G01N 33/574,</b> <b>C12Q 1/68</b>	<b>A2</b>	<b>(11) International Publication Number: WO 00/20587</b> <b>(43) International Publication Date: 13 April 2000 (13.04.00)</b>
<b>(21) International Application Number:</b> PCT/US99/22873 <b>(22) International Filing Date:</b> 4 October 1999 (04.10.99)  <b>(30) Priority Data:</b> 09/166,300                      5 October 1998 (05.10.98)            US 09/166,350                      5 October 1998 (05.10.98)            US  <b>(71) Applicant:</b> LUDWIG INSTITUTE FOR CANCER RE- SEARCH [CH/US]; 605 Third Avenue, New York, NY 10158 (US).  <b>(72) Inventors:</b> OBATA, Yuichi; Chikusa-Ku, Nagoya 464 (JP). GOUT, Ivan; 91 Riding House Street, London W1P 8BT (GB). TURECI, Ozlem; Innere Medizin 1, D-66421 Hom- burg/Saar (DE). SAHIN, Ugar; Innere Medizin 1, D-66421 Homburg/Saar (DE). PFREUNDSCHUH, Michael; Innere Medizin 1, D-66421 Homburg/Saar (DE). SCANLAN, Matthew, J.; 1275 York Avenue, New York, NY 10021 (US). STOCKERT, Elisabeth; 1275 York Avenue, New York, NY 10021 (US). CHEN, Yao-Tseng; 525 East 68th Street, New York, NY 10021 (US). OLD, Lloyd, J.; 605 Third Avenue, New York, NY 10158 (US). JAGER, Elke; Hohl 2-28, D-60488 Frankfurt am Main (DE). KNUTH, Alex; Hohl 2-28, D-60488 Frankfurt am Main (DE).		<b>(74) Agent:</b> VAN AMSTERDAM, John, R.; Wolf, Greenfield & Sacks, P.C., 600 Atlantic Avenue, Boston, MA 02210 (US).  <b>(81) Designated States:</b> AU, CA, CN, JP, KR, NZ, European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE).  <b>Published</b> <i>Without international search report and to be republished          upon receipt of that report.</i>
<b>(54) Title:</b> CANCER ASSOCIATED ANTIGENS AND USES THEREFOR  <b>(57) Abstract</b>  <p>Cancer associated antigens have been identified by autologous antibody screening of libraries of nucleic acids expressed in renal cancer cells using antisera from cancer patients. The invention relates to nucleic acids and encoded polypeptides which are cancer associated antigens expressed in patients afflicted with renal cancer. The invention provides, <i>inter alia</i>, isolated nucleic acid molecules, expression vectors containing those molecules and host cells transfected with those molecules. The invention also provides isolated proteins and peptides, antibodies to those proteins and peptides and cytotoxic T lymphocytes which recognize the proteins and peptides. Fragments of the foregoing including functional fragments and variants also are provided. Kits containing the foregoing molecules additionally are provided. The molecules provided by the invention can be used in the diagnosis, monitoring, research, or treatment of conditions characterized by the expression of one or more cancer associated antigens.</p>		

**FOR THE PURPOSES OF INFORMATION ONLY**

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece			TR	Turkey
BG	Bulgaria	HU	Hungary	ML	Mali	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MN	Mongolia	UA	Ukraine
BR	Brazil	IL	Israel	MR	Mauritania	UG	Uganda
BY	Belarus	IS	Iceland	MW	Malawi	US	United States of America
CA	Canada	IT	Italy	MX	Mexico	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NE	Niger	VN	Viet Nam
CG	Congo	KE	Kenya	NL	Netherlands	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NO	Norway	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	NZ	New Zealand		
CM	Cameroon			PL	Poland		
CN	China	KR	Republic of Korea	PT	Portugal		
CU	Cuba	KZ	Kazakstan	RO	Romania		
CZ	Czech Republic	LC	Saint Lucia	RU	Russian Federation		
DE	Germany	LI	Liechtenstein	SD	Sudan		
DK	Denmark	LK	Sri Lanka	SE	Sweden		
EE	Estonia	LR	Liberia	SG	Singapore		

## **CANCER ASSOCIATED ANTIGENS AND USES THEREFOR**

### **Field of the Invention**

The invention relates to nucleic acids and encoded polypeptides which are cancer associated antigens expressed in patients afflicted with a variety of cancers. The invention also relates to agents which bind the nucleic acids or polypeptides. The nucleic acid molecules, polypeptides coded for by such molecules and peptides derived therefrom, as well as related antibodies and cytolytic T lymphocytes, are useful, *inter alia*, in diagnostic and therapeutic contexts.

### **Background of the Invention**

The mechanism by which T cells recognize foreign materials has been implicated in cancer. A number of cytolytic T lymphocyte (CTL) clones directed against autologous melanoma antigens, testicular antigens, and melanocyte differentiation antigens have been described. In many instances, the antigens recognized by these clones have been characterized.

The use of autologous CTLs for identifying tumor antigens requires that the target cells which express the antigens can be cultured *in vitro* and that stable lines of autologous CTL clones which recognize the antigen-expressing cells can be isolated and propagated. While this approach has worked well for melanoma antigens, other tumor types, such as epithelial cancers including breast and colon cancer, have proved refractory to the approach.

More recently another approach to the problem has been described by Sahin et al. (*Proc. Natl. Acad. Sci. USA* 92:11810-11813, 1995). According to this approach, autologous antisera are used to identify immunogenic protein antigens expressed in cancer cells by screening expression libraries constructed from tumor cell cDNA. Antigen-encoding clones so identified have been found to have elicited an high-titer humoral immune response in the patients from which the antisera were obtained. Such a high-titer IgG response implies helper T cell recognition of the detected antigen. These tumor antigens can then be screened for the presence of MHC/HLA class I and class II motifs and reactivity with CTLs

Presently there is a need for additional cancer antigens for development of therapeutics and diagnosis applicable to a greater number of cancer patients having various cancers.

### Summary of the Invention

Autologous antibody screening has now been applied to renal cancer using antisera from cancer patients. Numerous cancer associated antigens have been identified. The invention provides, *inter alia*, isolated nucleic acid molecules, expression vectors containing those molecules and host cells transfected with those molecules. The invention also provides isolated proteins and peptides, antibodies to those proteins and peptides and CTLs which recognize the proteins and peptides. Fragments including functional fragments and variants of the foregoing also are provided. Kits containing the foregoing molecules additionally are provided. The foregoing can be used in the diagnosis, monitoring, research, or treatment of conditions characterized by the expression of one or more cancer associated antigens.

Prior to the present invention, only a handful of cancer associated genes had been identified in the past 20 years. The invention involves the surprising discovery of several genes, some previously known and some previously unknown, which are expressed in individuals who have cancer. These individuals all have serum antibodies against the proteins (or fragments thereof) encoded by these genes. Thus, abnormally expressed genes are recognized by the host's immune system and therefore can form a basis for diagnosis, monitoring and therapy.

The invention involves the use of a single material, a plurality of different materials and even large panels and combinations of materials. For example, a single gene, a single protein encoded by a gene, a single functional fragment thereof, a single antibody thereto, etc. can be used in methods and products of the invention. Likewise, pairs, groups and even panels of these materials and optionally other cancer associated antigen genes and/or gene products can be used for diagnosis, monitoring and therapy. The pairs, groups or panels can involve 2, 3, 4, 5 or more genes, gene products, fragments thereof or agents that recognize such materials. A plurality of such materials are not only useful in monitoring, typing, characterizing and diagnosing cells abnormally expressing such genes, but a plurality of such materials can be used therapeutically. An example of the use of a plurality of such materials for the prevention, delay of onset, amelioration, etc. of cancer cells, which express or will express such genes prophylactically or acutely. Any and all combinations of the genes, gene products, and materials which recognize the genes and gene products can be tested and identified for use according to the invention. It would be far too lengthy to recite all such combinations; those skilled in the art, particularly in view of the teaching contained herein,



will readily be able to determine which combinations are most appropriate for which circumstances.

As will be clear from the following discussion, the invention has *in vivo* and *in vitro* uses, including for therapeutic, diagnostic, monitoring and research purposes. One aspect of the invention is the ability to fingerprint a cell expressing a number of the genes identified according to the invention by, for example, quantifying the expression of such gene products. Such fingerprints will be characteristic, for example, of the stage of the cancer, the type of the cancer, or even the effect in animal models of a therapy on a cancer. Cells also can be screened to determine whether such cells abnormally express the genes identified according to the invention.

The invention, in one aspect, is a method of diagnosing a disorder characterized by expression of a cancer associated antigen precursor coded for by a nucleic acid molecule. The method involves the steps of contacting a biological sample isolated from a subject with an agent that specifically binds to the nucleic acid molecule, an expression product thereof, or a fragment of an expression product thereof complexed with an MHC, preferably an HLA, molecule, wherein the nucleic acid molecule is a NA Group 1 nucleic acid molecule, and determining the interaction between the agent and the nucleic acid molecule, the expression product or fragment of the expression product as a determination of the disorder.

In one embodiment the agent is selected from the group consisting of (a) a nucleic acid molecule comprising NA Group 1 nucleic acid molecules or a fragment thereof, (b) a nucleic acid molecule comprising NA Group 3 nucleic acid molecules or a fragment thereof, (c) a nucleic acid molecule comprising NA Group 5 nucleic acid molecules or a fragment thereof, (d) an antibody that binds to an expression product, or a fragment thereof, of NA group 1 nucleic acids, (e) an antibody that binds to an expression product, or a fragment thereof, of NA group 3 nucleic acids, (f) an antibody that binds to an expression product, or a fragment thereof, of NA group 5 nucleic acids, (g) an agent that binds to a complex of an MHC, preferably HLA, molecule and a fragment of an expression product of a NA Group 1 nucleic acid, (h) an agent that binds to a complex of an MHC, preferably HLA, molecule and a fragment of an expression product of a NA group 3 nucleic acid, and (i) an agent that binds to a complex of an MHC, preferably HLA, molecule and a fragment of an expression product of a NA Group 5 nucleic acid.

The disorder may be characterized by expression of a plurality of cancer associated

antigen precursors. Thus the methods of diagnosis may include use of a plurality of agents, each of which is specific for a different human cancer associated antigen precursor (including at least one of the cancer associated antigen precursors disclosed herein), and wherein said plurality of agents is at least 2, at least 3, at least 4, at least 5, at least 6, at least 7, at least 8, at least 9 or at least 10 such agents.

In each of the above embodiments the agent may be specific for a human cancer associated antigen precursor that is selected from the group consisting of breast, gastric, lung, prostate, renal, colon, thyroid, Hodgkin's disease, and hepatocarcinoma cancer associated antigen precursors.

In another aspect the invention is a method for determining regression, progression or onset of a condition characterized by expression of abnormal levels of a protein encoded by a nucleic acid molecule that is a NA Group 1 molecule. The method involves the steps of monitoring a sample, from a subject who has or is suspected of having the condition, for a parameter selected from the group consisting of (i) the protein, (ii) a peptide derived from the protein, (iii) an antibody which selectively binds the protein or peptide, and (iv) cytolytic T cells specific for a complex of the peptide derived from the protein and an MHC molecule, as a determination of regression, progression or onset of said condition. In one embodiment the sample is a body fluid, a body effusion or a tissue.

In another embodiment the step of monitoring comprises contacting the sample with a detectable agent selected from the group consisting of (a) an antibody which selectively binds the protein of (i), or the peptide of (ii), (b) a protein or peptide which binds the antibody of (iii), and (c) a cell which presents the complex of the peptide and MHC molecule of (iv). In a preferred embodiment the antibody, the protein, the peptide or the cell is labeled with a radioactive label or an enzyme. The sample in a preferred embodiment is assayed for the peptide.

According to another embodiment the nucleic acid molecule is one of the following: a NA Group 3 molecule or a NA Group 5 molecule. In yet another embodiment the protein is a plurality of proteins, the parameter is a plurality of parameters, each of the plurality of parameters being specific for a different of the plurality of proteins. In certain embodiments the protein is a plurality of proteins, at least one of which is kinectin, the remainder of which are non-kinectin cancer associated proteins, and wherein the parameter is a plurality of parameters, each of the plurality of parameters being specific for a different of the plurality of

proteins.

The invention in another aspect is a pharmaceutical preparation for a human subject. The pharmaceutical preparation includes an agent which when administered to the subject enriches selectively the presence of complexes of an HLA molecule and a human cancer associated antigen, and a pharmaceutically acceptable carrier, wherein the human cancer associated antigen is a fragment of a human cancer associated antigen precursor encoded by a nucleic acid molecule which comprises a NA Group 1 molecule. In one embodiment the nucleic acid molecule is a NA Group 3 nucleic acid molecule.

The agent in one embodiment comprises a plurality of agents, each of which enriches selectively in the subject complexes of an HLA molecule and a different human cancer associated antigen. Preferably the plurality is at least two, at least three, at least four or at least 5 different such agents.

In certain embodiments, the agent comprises a plurality of agents, at least one of which is kinectin, the remainder of which are non-kinectin cancer associated proteins, and each of which enriches selectively in the subject complexes of an HLA molecule and a different human cancer associated antigen.

In another embodiment the agent is selected from the group consisting of (1) an isolated polypeptide comprising the human cancer associated antigen, or a functional variant thereof, (2) an isolated nucleic acid operably linked to a promoter for expressing the isolated polypeptide, or functional variant thereof, (3) a host cell expressing the isolated polypeptide, or functional variant thereof, and (4) isolated complexes of the polypeptide, or functional variants thereof, and an HLA molecule.

The agent may be a cell expressing an isolated polypeptide. In one embodiment the agent is a cell expressing an isolated polypeptide comprising the human cancer associated antigen or a functional variant thereof. In another embodiment the agent is a cell expressing an isolated polypeptide comprising the human cancer associated antigen or a functional variant thereof, and wherein the cell expresses an HLA molecule that binds the polypeptide. The cell can express one or both of the polypeptide and HLA molecule recombinantly. In preferred embodiments the cell is nonproliferative. In other preferred embodiments, the isolated polypeptide is or includes a kinectin polypeptide. In yet another embodiment the agent is at least two, at least three, at least four or at least five different polypeptides, each representing a different human cancer associated antigen or functional variant thereof.

The agent in one embodiment is a PP Group 2 polypeptide. In other embodiments the agent is a PP Group 3 polypeptide or a PP Group 4 polypeptide.

In an embodiment each of the pharmaceutical preparations described herein also includes an adjuvant.

According to another aspect the invention, a composition is provided which includes an isolated agent that binds selectively a PP Group 1 polypeptide. In separate embodiments the agent binds selectively to a polypeptide selected from the following: a PP Group 2 polypeptide, a PP Group 3 polypeptide, a PP Group 4 polypeptide, and a PP Group 5 polypeptide. In other embodiments, the agent is a plurality of different agents that bind selectively at least two, at least three, at least four, or at least five different such polypeptides. In each of the above described embodiments the agent may be an antibody. In a preferred embodiment, at least one of polypeptides is kinectin or a fragment thereof.

In another aspect the invention is a composition of matter composed of a conjugate of the agent of the above-described compositions of the invention and a therapeutic or diagnostic agent. Preferably the conjugate is of the agent and a therapeutic or diagnostic that is an antineoplastic.

The invention in another aspect is a pharmaceutical composition which includes an isolated nucleic acid molecule selected from the group consisting of: (1) NA Group 1 molecules, and (2) NA Group 2 molecules, and a pharmaceutically acceptable carrier. In one embodiment the isolated nucleic acid molecule comprises a NA Group 3 or NA Group 4 molecule. In another embodiment the isolated nucleic acid molecule comprises at least two isolated nucleic acid molecules coding for two different polypeptides, each polypeptide comprising a different cancer associated antigen. In preferred embodiments, at least one of the polypeptides is a kinectin polypeptide.

Preferably the pharmaceutical composition also includes an expression vector with a promoter operably linked to the isolated nucleic acid molecule. In another embodiment the pharmaceutical composition also includes a host cell recombinantly expressing the isolated nucleic acid molecule.

According to another aspect of the invention a pharmaceutical composition is provided. The pharmaceutical composition includes an isolated polypeptide comprising a PP Group 1 or a PP Group 2 polypeptide, and a pharmaceutically acceptable carrier. In one embodiment the isolated polypeptide comprises a PP Group 3 or a PP Group 4 polypeptide.

In another embodiment the isolated polypeptide comprises at least two different polypeptides, each comprising a different cancer associated antigen at least one of which is encoded by a NA group 1 molecule as disclosed herein. In separate embodiments the isolated polypeptides are selected from the following: PP Group 3 polypeptides or HLA binding fragments thereof and PP Group 5 polypeptides or HLA binding fragments thereof.

In an embodiment each of the pharmaceutical compositions described herein also includes an adjuvant.

Another aspect the invention is an isolated nucleic acid molecule comprising a NA Group 3 molecule. Another aspect the invention is an isolated nucleic acid molecule comprising a NA Group 4 molecule.

The invention in another aspect is an isolated nucleic acid molecule selected from the group consisting of (a) a fragment of a nucleic acid selected from the group of nucleic acid molecules consisting of SEQ ID Nos numbered below and comprising all nucleic acid sequences among SEQ ID NOs 1-11 and 22-46, of sufficient length to represent a sequence unique within the human genome, and identifying a nucleic acid encoding a human cancer associated antigen precursor, (b) complements of (a), provided that the fragment includes a sequence of contiguous nucleotides which is not identical to any sequence selected from the sequence group consisting of (1) sequences having the GenBank accession numbers of Table 1, (2) complements of (1), and (3) fragments of (1) and (2).

In one embodiment the sequence of contiguous nucleotides is selected from the group consisting of: (1) at least two contiguous nucleotides nonidentical to the sequences in Table 1, (2) at least three contiguous nucleotides nonidentical to the sequences in Table 1, (3) at least four contiguous nucleotides nonidentical to the sequences in Table 1, (4) at least five contiguous nucleotides nonidentical to the sequences in Table 1, (5) at least six contiguous nucleotides nonidentical to the sequences in Table 1, or (6) at least seven contiguous nucleotides nonidentical to the sequences in Table 1.

In another embodiment the fragment has a size selected from the group consisting of at least: 8 nucleotides, 10 nucleotides, 12 nucleotides, 14 nucleotides, 16 nucleotides, 18 nucleotides, 20, nucleotides, 22 nucleotides, 24 nucleotides, 26 nucleotides, 28 nucleotides, 30 nucleotides, 50 nucleotides, 75 nucleotides, 100 nucleotides, 200 nucleotides, 1000 nucleotides and every integer length therebetween.

In yet another embodiment the molecule encodes a polypeptide which, or a fragment

of which, binds a human HLA receptor or a human antibody.

Another aspect of the invention is an expression vector comprising an isolated nucleic acid molecule of the invention described above operably linked to a promoter.

According to one aspect the invention is an expression vector comprising a nucleic acid operably linked to a promoter, wherein the nucleic acid is a NA Group 1 or Group 2 molecule. In another aspect the invention is an expression vector comprising a NA Group 1 or Group 2 molecule and a nucleic acid encoding an MHC, preferably HLA, molecule.

In yet another aspect the invention is a host cell transformed or transfected with an expression vector of the invention described above.

In another aspect the invention is a host cell transformed or transfected with an expression vector comprising an isolated nucleic acid molecule of the invention described above operably linked to a promoter, or an expression vector comprising a nucleic acid operably linked to a promoter, wherein the nucleic acid is a NA Group 1 or 2 molecule and further comprising a nucleic acid encoding HLA.

According to another aspect of the invention an isolated polypeptide encoded by the isolated nucleic acid molecules the invention, described above, is provided. These include PP Group 1-5 polypeptides. The invention also includes a fragment of the polypeptide which is immunogenic. In one embodiment the fragment, or a portion of the fragment, binds HLA or a human antibody. In still another aspect the invention provides as isolated polypeptide comprising a fragment of a kinectin polypeptide which is immunogenic.

The invention includes in another aspect an isolated fragment of a human cancer associated antigen precursor which, or portion of which, binds HLA or a human antibody, wherein the precursor is encoded by a nucleic acid molecule that is a NA Group 1 molecule. In one embodiment the fragment is part of a complex with HLA. In another embodiment the fragment is between 8 and 12 amino acids in length. In another embodiment the invention includes an isolated polypeptide comprising a fragment of the polypeptide of sufficient length to represent a sequence unique within the human genome and identifying a polypeptide that is a human cancer associated antigen precursor.

According to another aspect of the invention a kit for detecting the presence of the expression of a cancer associated antigen precursor is provided. The kit includes a pair of isolated nucleic acid molecules each of which consists essentially of a molecule selected from the group consisting of (a) a 12-32 nucleotide contiguous segment of the nucleotide sequence

of any of the NA Group 1 molecules and (b) complements of ("a"), wherein the contiguous segments are nonoverlapping. In one embodiment the pair of isolated nucleic acid molecules is constructed and arranged to selectively amplify an isolated nucleic acid molecule that is a NA Group 3 molecule. Preferably, the pair amplifies a human NA Group 3 molecule.

According to another aspect of the invention a method for treating a subject with a disorder characterized by expression of a human cancer associated antigen precursor is provided. The method includes the step of administering to the subject an amount of an agent, which enriches selectively in the subject the presence of complexes of an HLA molecule and a human cancer associated antigen, effective to ameliorate the disorder, wherein the human cancer associated antigen is a fragment of a human cancer associated antigen precursor encoded by a nucleic acid molecule selected from the group consisting of (a) a nucleic acid molecule comprising NA group 1 nucleic acid molecules, (b) a nucleic acid molecule comprising NA group 3 nucleic acid molecules, (c) a nucleic acid molecule comprising NA group 5 nucleic acid molecules.

In one embodiment the disorder is characterized by expression of a plurality of human cancer associated antigen precursors and wherein the agent is a plurality of agents, each of which enriches selectively in the subject the presence of complexes of an HLA molecule and a different human cancer associated antigen. Preferably the plurality is at least 2, at least 3, at least 4, or at least 5 such agents. In a preferred embodiment, at least one of the human cancer associated antigens is kinectin or a fragment thereof.

In another embodiment the agent is an isolated polypeptide selected from the group consisting of PP Group 1, PP Group 2, PP Group 3, PP Group 4, and PP group 5 polypeptides.

In yet another embodiment the disorder is cancer.

According to another aspect the invention is a method for treating a subject having a condition characterized by expression of a cancer associated antigen precursor in cells of the subject. The method includes the steps of (i) removing an immunoreactive cell containing sample from the subject, (ii) contacting the immunoreactive cell containing sample to the host cell under conditions favoring production of cytolytic T cells against a human cancer associated antigen which is a fragment of the precursor, (iii) introducing the cytolytic T cells to the subject in an amount effective to lyse cells which express the human cancer associated antigen, wherein the host cell is transformed or transfected with an expression vector comprising an isolated nucleic acid molecule operably linked to a promoter, the isolated

nucleic acid molecule being selected from the group of nucleic acid molecules consisting of NA Group 1, NA Group 2, NA Group 3, NA Group 4, NA Group 5.

In one embodiment the host cell recombinantly expresses an HLA molecule which binds the human cancer associated antigen. In another embodiment the host cell endogenously expresses an HLA molecule which binds the human cancer associated antigen.

The invention includes in another aspect a method for treating a subject having a condition characterized by expression of a cancer associated antigen precursor in cells of the subject. The method includes the steps of (i) identifying a nucleic acid molecule expressed by the cells associated with said condition, wherein said nucleic acid molecule is a NA Group 1 molecule (ii) transfecting a host cell with a nucleic acid selected from the group consisting of (a) the nucleic acid molecule identified, (b) a fragment of the nucleic acid identified which includes a segment coding for a cancer associated antigen, (c) deletions, substitutions or additions to (a) or (b), and (d) degenerates of (a), (b), or (c); (iii) culturing said transfected host cells to express the transfected nucleic acid molecule, and; (iv) introducing an amount of said host cells or an extract thereof to the subject effective to increase an immune response against the cells of the subject associated with the condition. Preferably, the antigen is a human antigen and the subject is a human. In certain preferred embodiments the nucleic acid molecule is a kinectin nucleic acid molecule.

In one embodiment the method also includes the step of (a) identifying an MHC molecule which presents a portion of an expression product of the nucleic acid molecule, wherein the host cell expresses the same MHC molecule as identified in (a) and wherein the host cell presents an MHC binding portion of the expression product of the nucleic acid molecule.

In another embodiment the method also includes the step of treating the host cells to render them non-proliferative.

In yet another embodiment the immune response comprises a B-cell response or a T cell response. Preferably the response is a T-cell response which comprises generation of cytolytic T-cells specific for the host cells presenting the portion of the expression product of the nucleic acid molecule or cells of the subject expressing the human cancer associated antigen.

In another embodiment the nucleic acid molecule is a NA Group 3 molecule.

Another aspect of the invention is a method for treating or diagnosing or monitoring a



subject having a condition characterized by expression of an abnormal amount of a protein encoded by a nucleic acid molecule that is a NA Group 1 molecule. The method includes the step of administering to the subject an antibody which specifically binds to the protein or a peptide derived therefrom, the antibody being coupled to a therapeutically useful agent, in an amount effective to treat the condition.

In one embodiment the antibody is a monoclonal antibody. Preferably the monoclonal antibody is a chimeric antibody or a humanized antibody.

In another aspect the invention is a method for treating a condition characterized by expression in a subject of abnormal amounts of a protein encoded by a nucleic acid molecule that is a NA Group 1 nucleic acid molecule. The method involves the step of administering to a subject at least one of the pharmaceutical compositions of the invention described above in an amount effective to prevent, delay the onset of, or inhibit the condition in the subject. In one embodiment the condition is cancer. In another embodiment the method includes the step of first identifying that the subject expresses in a tissue abnormal amounts of the protein.

The invention in another aspect is a method for treating a subject having a condition characterized by expression of abnormal amounts of a protein encoded by a nucleic acid molecule that is a NA Group 1 nucleic acid molecule. The method includes the steps of (i) identifying cells from the subject which express abnormal amounts of the protein; (ii) isolating a sample of the cells; (iii) cultivating the cells, and (iv) introducing the cells to the subject in an amount effective to provoke an immune response against the cells.

In one embodiment the method includes the step of rendering the cells non-proliferative, prior to introducing them to the subject.

In another aspect the invention is a method for treating a pathological cell condition characterized by abnormal expression of a protein encoded by a nucleic acid molecule that is a NA Group 1 nucleic acid molecule. The method includes the step of administering to a subject in need thereof an effective amount of an agent which inhibits the expression or activity of the protein.

In one embodiment the agent is an inhibiting antibody which selectively binds to the protein and wherein the antibody is a monoclonal antibody, a chimeric antibody, a humanized antibody or a fragment thereof. In another embodiment the agent is an antisense nucleic acid molecule which selectively binds to the nucleic acid molecule which encodes the protein. In yet another important embodiment the nucleic acid molecule is a NA Group 3 nucleic acid

molecule. In other preferred embodiments, the nucleic acid molecule is a kinectin nucleic acid molecule.

The invention includes in another aspect a composition of matter useful in stimulating an immune response to a plurality of proteins encoded by nucleic acid molecules that are NA Group 1 molecules. The composition is a plurality of peptides derived from the amino acid sequences of the proteins, wherein the peptides bind to one or more MHC molecules presented on the surface of the cells which express an abnormal amount of the protein. In preferred embodiments, at least one of the proteins is kinectin.

In one embodiment at least a portion of the plurality of peptides bind to MHC molecules and elicit a cytolytic response thereto. In another embodiment the composition of matter includes an adjuvant. In another embodiment the adjuvant is a saponin, GM-CSF, or an interleukin. In still another embodiment, the compositions also includes at least one peptide useful in stimulating an immune response to at least one protein which is not encoded by nucleic acid molecules that are NA Group 1 molecules, wherein the at least one peptide binds to one or more MHC molecules.

According to another aspect the invention is an isolated antibody which selectively binds to a complex of: (i) a peptide derived from a protein encoded by a nucleic acid molecule that is a NA Group 1 molecule and (ii) and an MHC molecule to which binds the peptide to form the complex, wherein the isolated antibody does not bind to (i) or (ii) alone.

In one embodiment the antibody is a monoclonal antibody, a chimeric antibody, a humanized antibody or a fragment thereof.

The invention also involves the use of the genes, gene products, fragments thereof, agents which bind thereto, and so on in the preparation of medicaments. A particular medicament is for treating cancer and a more particular medicament is for treating breast cancer, lung cancer, renal cancer, colon cancer, prostate cancer or gastric cancer.

### **Detailed Description of the Invention**

In the above summary and in the ensuing description, lists of sequences are provided. The lists are meant to embrace each single sequence separately, two or more sequences together where they form a part of the same gene, any combination of two or more sequences which relate to different genes, including and up to the total number on the list, as if each and every combination were separately and specifically enumerated. Likewise, when mentioning

fragment size, it is intended that a range embrace the smallest fragment mentioned to the full-length of the sequence (less one nucleotide or amino acid so that it is a fragment), each and every fragment length intended as if specifically enumerated. Thus, if a fragment could be between 10 and 15 in length, it is explicitly meant to mean 10, 11, 12, 13, 14, or 15 in length.

The summary and the claims mention antigen precursors and antigens. As used in the summary and in the claims, a precursor is substantially the full-length protein encoded by the coding region of the isolated DNA and the antigen is a peptide which complexes with MHC, preferably HLA, and which participates in the immune response as part of that complex. Such antigens are typically 9 amino acids long, although this may vary slightly.

As used herein, a subject is a human, non-human primate, cow, horse, pig, sheep, goat, dog, cat or rodent. In all embodiments human cancer antigens and human subjects are preferred.

The present invention in one aspect involves the cloning of cDNAs encoding human cancer associated antigen precursors using autologous antisera of subjects having renal cancer. The sequences of the clones representing genes identified according to the methods described herein are presented in the attached Sequence Listing. Of the foregoing, it can be seen that some of the clones are considered completely novel as no nucleotide or amino acid homologies to coding regions were found in the databases searched. Other clones are novel but have some homology to sequences deposited in databases (mainly EST sequences). Nevertheless, the entire gene sequence was not previously known. In some cases no function was suspected and in other cases, even if a function was suspected, it was not known that the gene was associated with cancer. In all cases, it was not known or suspected that the gene encoded a cancer antigen which reacted with antibody from autologous sera. Analysis of the clone sequences by comparison to nucleic acid and protein databases determined that still other of the clones surprisingly are closely related to other previously-cloned genes. The sequences of these related genes is also presented in the Sequence Listing. The nature of the foregoing genes as encoding antigens recognized by the immune systems of cancer patients is, of course, unexpected.

The invention thus involves in one aspect cancer associated antigen polypeptides, genes encoding those polypeptides, functional modifications and variants of the foregoing, useful fragments of the foregoing, as well as diagnostics and therapeutics relating thereto.

Homologs and alleles of the cancer associated antigen nucleic acids of the invention

can be identified by conventional techniques. Thus, an aspect of the invention is those nucleic acid sequences which code for cancer associated antigen precursors. Because this application contains so many sequences, the following chart is provided to identify the various groups of sequences discussed in the claims and in the summary:

### Nucleic Acid Sequences

NA Group 1. (a) nucleic acid molecules which hybridize under stringent conditions to a molecule consisting of a nucleic acid sequence selected from the group consisting of nucleic acid sequences among SEQ ID NOs: 1-11 and 22-46 and which code for a cancer associated antigen precursor,

(b) deletions, additions and substitutions which code for a respective cancer associated antigen precursor,

(c) nucleic acid molecules that differ from the nucleic acid molecules of (a) or (b) in codon sequence due to the degeneracy of the genetic code, and

(d) complements of (a), (b) or (c).

NA Group 2. Fragments of NA Group 1, which codes for a polypeptide which, or a portion of which, binds an MHC molecule to form a complex recognized by a an autologous antibody or lymphocyte.

NA Group 3. The subset of NA Group 1 where the nucleotide sequence is selected from the group consisting of:

(a) previously unknown human nucleic acids coding for a human cancer associated antigen precursor (i.e. nucleic acid sequences among SEQ ID NOs: 1-11),

(b) deletions, additions and substitutions which code for a respective human cancer associated antigen precursor,

(c) nucleic acid molecules that differ from the nucleic acid molecules of (a) or (b) in codon sequence due to the degeneracy of the genetic code, and

(d) complements of (a), (b) or (c).

NA Group 4. Fragments of NA Group 3, which code for a polypeptide which, or a portion of which, binds to an MHC molecule to form a complex recognized by an autologous antibody or lymphocyte.

NA Group 5. A subset of NA Group 1, comprising human cancer associated antigens that react with allogeneic cancer antisera.

#### Polypeptide Sequences

- PP Group 1. Polypeptides encoded by NA Group 1.
- PP Group 2. Polypeptides encoded by NA Group 2
- PP Group 3. Polypeptides encoded by NA Group 3.
- PP Group 4. Polypeptides encoded by NA Group 4.
- PP Group 5. Polypeptides encoded by NA Group 5.

The term "stringent conditions" as used herein refers to parameters with which the art is familiar. Nucleic acid hybridization parameters may be found in references which compile such methods, e.g. *Molecular Cloning: A Laboratory Manual*, J. Sambrook, et al., eds., Second Edition, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York, 1989, or *Current Protocols in Molecular Biology*, F.M. Ausubel, et al., eds., John Wiley & Sons, Inc., New York. More specifically, stringent conditions, as used herein, refers, for example, to hybridization at 65°C in hybridization buffer (3.5 x SSC, 0.02% Ficoll, 0.02% polyvinyl pyrrolidone, 0.02% Bovine Serum Albumin, 2.5mM NaH<sub>2</sub>PO<sub>4</sub>(pH7), 0.5% SDS, 2mM EDTA). SSC is 0.15M sodium chloride/0.15M sodium citrate, pH7; SDS is sodium dodecyl sulphate; and EDTA is ethylenediaminetetracetic acid. After hybridization, the membrane upon which the DNA is transferred is washed, for example, in 2 x SSC at room temperature and then at 0.1 - 0.5 x SSC/0.1 x SDS at temperatures up to 68°C.

There are other conditions, reagents, and so forth which can be used, which result in a similar degree of stringency. The skilled artisan will be familiar with such conditions, and thus they are not given here. It will be understood, however, that the skilled artisan will be able to manipulate the conditions in a manner to permit the clear identification of homologs and alleles of cancer associated antigen nucleic acids of the invention (e.g., by using lower stringency conditions). The skilled artisan also is familiar with the methodology for screening

cells and libraries for expression of such molecules which then are routinely isolated, followed by isolation of the pertinent nucleic acid molecule and sequencing.

In general homologs and alleles typically will share at least 75% nucleotide identity and/or at least 90% amino acid identity to the sequences of cancer associated antigen nucleic acid and polypeptides, respectively, in some instances will share at least 90% nucleotide identity and/or at least 95% amino acid identity and in still other instances will share at least 95% nucleotide identity and/or at least 99% amino acid identity. The homology can be calculated using various, publicly available software tools developed by NCBI (Bethesda, Maryland) that can be obtained through the internet (<ftp://ncbi.nlm.nih.gov/pub/>). Exemplary tools include the BLAST system available at <http://www.ncbi.nlm.nih.gov>, using default settings. Pairwise and ClustalW alignments (BLOSUM30 matrix setting) as well as Kyte-Doolittle hydropathic analysis can be obtained using the MacVector sequence analysis software (Oxford Molecular Group). Watson-Crick complements of the foregoing nucleic acids also are embraced by the invention.

In screening for cancer associated antigen genes, a Southern blot may be performed using the foregoing conditions, together with a radioactive probe. After washing the membrane to which the DNA is finally transferred, the membrane can be placed against X-ray film to detect the radioactive signal. In screening for the expression of cancer associated antigen nucleic acids, Northern blot hybridizations using the foregoing conditions (see also the Examples) can be performed on samples taken from breast cancer patients or subjects suspected of having a condition characterized by expression of breast cancer associated antigen genes. Amplification protocols such as polymerase chain reaction using primers which hybridize to the sequences presented also can be used for detection of the cancer associated antigen genes or expression thereof.

The renal cancer associated genes correspond to SEQ ID NOs. 1-11 and 22-35. Kinectin cancer associated sequences correspond to SEQ ID Nos:36-46. The preferred cancer associated antigens for the methods of diagnosis disclosed herein are those which were found to react with allogeneic cancer antisera (i.e. NA Group 5). Encoded polypeptides (e.g., proteins), peptides and antisera thereto are also preferred for diagnosis.

The invention also includes degenerate nucleic acids which include alternative codons to those present in the native materials. For example, serine residues are encoded by the codons TCA, AGT, TCC, TCG, TCT and AGC. Each of the six codons is equivalent for the

purposes of encoding a serine residue. Thus, it will be apparent to one of ordinary skill in the art that any of the serine-encoding nucleotide triplets may be employed to direct the protein synthesis apparatus, *in vitro* or *in vivo*, to incorporate a serine residue into an elongating breast cancer associated antigen polypeptide. Similarly, nucleotide sequence triplets which encode other amino acid residues include, but are not limited to: CCA, CCC, CCG and CCT (proline codons); CGA, CGC, CGG, CGT, AGA and AGG (arginine codons); ACA, ACC, ACG and ACT (threonine codons); AAC and AAT (asparagine codons); and ATA, ATC and ATT (isoleucine codons). Other amino acid residues may be encoded similarly by multiple nucleotide sequences. Thus, the invention embraces degenerate nucleic acids that differ from the biologically isolated nucleic acids in codon sequence due to the degeneracy of the genetic code.

The invention also provides modified nucleic acid molecules which include additions, substitutions and deletions of one or more nucleotides. In preferred embodiments, these modified nucleic acid molecules and/or the polypeptides they encode retain at least one activity or function of the unmodified nucleic acid molecule and/or the polypeptides, such as antigenicity, enzymatic activity, receptor binding, formation of complexes by binding of peptides by MHC class I and class II molecules, etc. In certain embodiments, the modified nucleic acid molecules encode modified polypeptides, preferably polypeptides having conservative amino acid substitutions as are described elsewhere herein. The modified nucleic acid molecules are structurally related to the unmodified nucleic acid molecules and in preferred embodiments are sufficiently structurally related to the unmodified nucleic acid molecules so that the modified and unmodified nucleic acid molecules hybridize under stringent conditions known to one of skill in the art.

For example, modified nucleic acid molecules which encode polypeptides having single amino acid changes can be prepared. Each of these nucleic acid molecules can have one, two or three nucleotide substitutions exclusive of nucleotide changes corresponding to the degeneracy of the genetic code as described herein. Likewise, modified nucleic acid molecules which encode polypeptides having two amino acid changes can be prepared which have, e.g., 2-6 nucleotide changes. Numerous modified nucleic acid molecules like these will be readily envisioned by one of skill in the art, including for example, substitutions of nucleotides in codons encoding amino acids 2 and 3, 2 and 4, 2 and 5, 2 and 6, and so on. In the foregoing example, each combination of two amino acids is included in the set of

modified nucleic acid molecules, as well as all nucleotide substitutions which code for the amino acid substitutions. Additional nucleic acid molecules that encode polypeptides having additional substitutions (i.e., 3 or more), additions or deletions (e.g., by introduction of a stop codon or a splice site(s)) also can be prepared and are embraced by the invention as readily envisioned by one of ordinary skill in the art. Any of the foregoing nucleic acids or polypeptides can be tested by routine experimentation for retention of structural relation or activity to the nucleic acids and/or polypeptides disclosed herein.

The invention also provides isolated unique fragments of cancer associated antigen nucleic acid sequences or complements thereof. A unique fragment is one that is a 'signature' for the larger nucleic acid. It, for example, is long enough to assure that its precise sequence is not found in molecules within the human genome outside of the cancer associated antigen nucleic acids defined above (and human alleles). Those of ordinary skill in the art may apply no more than routine procedures to determine if a fragment is unique within the human genome. Unique fragments, however, exclude fragments completely composed of the nucleotide sequences of any of GenBank accession numbers listed in Table 1 or other previously published sequences as of the filing date of the priority documents for sequences listed in a respective priority document or the filing date of this application for sequences listed for the first time in this application which overlap the sequences of the invention.

A fragment which is completely composed of the sequence described in the foregoing GenBank deposits is one which does not include any of the nucleotides unique to the sequences of the invention. Thus, a unique fragment must contain a nucleotide sequence other than the exact sequence of those in GenBank or fragments thereof. The difference may be an addition, deletion or substitution with respect to the GenBank sequence or it may be a sequence wholly separate from the GenBank sequence.

Unique fragments can be used as probes in Southern and Northern blot assays to identify such nucleic acids, or can be used in amplification assays such as those employing PCR. As known to those skilled in the art, large probes such as 200, 250, 300 or more nucleotides are preferred for certain uses such as Southern and Northern blots, while smaller fragments will be preferred for uses such as PCR. Unique fragments also can be used to produce fusion proteins for generating antibodies or determining binding of the polypeptide fragments, or for generating immunoassay components. Likewise, unique fragments can be employed to produce nonfused fragments of the cancer associated antigen polypeptides,



useful, for example, in the preparation of antibodies, and in immunoassays. Unique fragments further can be used as antisense molecules to inhibit the expression of cancer associated antigen nucleic acids and polypeptides, particularly for therapeutic purposes as described in greater detail below.

As will be recognized by those skilled in the art, the size of the unique fragment will depend upon its conservancy in the genetic code. Thus, some regions of cancer associated antigen sequences and complements thereof will require longer segments to be unique while others will require only short segments, typically between 12 and 32 nucleotides (e.g. 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31 and 32 or more bases long, up to the entire length of the disclosed sequence. As mentioned above, this disclosure intends to embrace each and every fragment of each sequence, beginning at the first nucleotide, the second nucleotide and so on, up to 8 nucleotides short of the end, and ending anywhere from nucleotide number 8, 9, 10 and so on for each sequence, up to the very last nucleotide (provided the sequence is unique as described above).

Virtually any segment of the polypeptide coding region of novel cancer associated antigen nucleic acids, or complements thereof, that is 18 or more nucleotides in length will be unique. Those skilled in the art are well versed in methods for selecting such sequences, typically on the basis of the ability of the unique fragment to selectively distinguish the sequence of interest from other sequences in the human genome of the fragment to those on known databases typically is all that is necessary, although *in vitro* confirmatory hybridization and sequencing analysis may be performed.

Especially preferred include nucleic acids encoding a series of epitopes, known as "polytopes". The epitopes can be arranged in sequential or overlapping fashion (*see, e.g.*, Thomson et al., *Proc. Natl. Acad. Sci. USA* 92:5845-5849, 1995; Gilbert et al., *Nature Biotechnol.* 15:1280-1284, 1997), with or without the natural flanking sequences, and can be separated by unrelated linker sequences if desired. The polytope is processed to generated individual epitopes which are recognized by the immune system for generation of immune responses.

Thus, for example, peptides derived from a polypeptide having an amino acid sequence encoded by one of the nucleic acid disclosed herein, and which are presented by MHC molecules and recognized by CTL or T helper lymphocytes, can be combined with peptides from one or more other cancer associated antigens (e.g. by preparation of hybrid

nucleic acids or polypeptides) to form "polytopes". The two or more peptides (or nucleic acids encoding the peptides) can be selected from those described herein, or they can include one or more peptides of previously known cancer associated antigens. Exemplary cancer associated peptide antigens that can be administered to induce or enhance an immune response are derived from tumor associated genes and encoded proteins including MAGE-A1, MAGE-A2, MAGE-A3, MAGE-A4, MAGE-A5, MAGE-A6, MAGE-A7, MAGE-A8, MAGE-A9, MAGE-A10, MAGE-A11, MAGE-A12, MAGE-13, GAGE-1, GAGE-2, GAGE-3, GAGE-4, GAGE-5, GAGE-6, GAGE-7, GAGE-8, BAGE-1, RAGE-1, LB33/MUM-1, PRAME, NAG, MAGE-B2, MAGE-B3, MAGE-B4, tyrosinase, brain glycogen phosphorylase, Melan-A, MAGE-C1 MAGE-C2, NY-ESO-1, LAGE-1, SSX-1, SSX-2 (HOM-MEL-40) SSX-4, SSX-5, SCP-1 and CT-7. See, for example, PCT application publication no. WO96/10577. Other examples will be known to one of ordinary skill in the art (for example, see Coulie, *Stem Cells* 13:393-403, 1995), and can be used in the invention in a like manner as those disclosed herein. One of ordinary skill in the art can prepare polypeptides comprising one or more peptides and one or more of the foregoing cancer associated peptides, or nucleic acids encoding such polypeptides, according to standard procedures of molecular biology.

Thus polytopes are groups of two or more potentially immunogenic or immune response stimulating peptides which can be joined together in various arrangements (e.g. concatenated, overlapping). The polytope (or nucleic acid encoding the polytope) can be administered in a standard immunization protocol, e.g. to animals, to test the effectiveness of the polytope in stimulating, enhancing and/or provoking an immune response.

The peptides can be joined together directly or via the use of flanking sequences to form polytopes, and the use of polytopes as vaccines is well known in the art (see, e.g., Thomson et al., *Proc. Acad. Natl. Acad. Sci USA* 92(13):5845-5849, 1995; Gilbert et al., *Nature Biotechnol.* 15(12):1280-1284, 1997; Thomson et al., *J. Immunol.* 157(2):822-826, 1996; Tam et al., *J. Exp. Med.* 171(1):299-306, 1990). For example, Tam showed that polytopes consisting of both MHC class I and class II binding epitopes successfully generated antibody and protective immunity in a mouse model. Tam also demonstrated that polytopes comprising "strings" of epitopes are processed to yield individual epitopes which are presented by MHC molecules and recognized by CTLs. Thus polytopes containing various numbers and combinations of epitopes can be prepared and tested for recognition by CTLs

and for efficacy in increasing an immune response.

It is known that tumors express a set of tumor antigens, of which only certain subsets may be expressed in the tumor of any given patient. Polytopes can be prepared which correspond to the different combination of epitopes representing the subset of tumor rejection antigens expressed in a particular patient. Polytopes also can be prepared to reflect a broader spectrum of tumor rejection antigens known to be expressed by a tumor type. Polytopes can be introduced to a patient in need of such treatment as polypeptide structures, or via the use of nucleic acid delivery systems known in the art (see, e.g., Allsopp et al., *Eur. J. Immunol.* 26(8):1951-1959, 1996). Adenovirus, pox virus, Ty-virus like particles, adeno-associated virus, plasmids, bacteria, etc. can be used in such delivery. One can test the polytope delivery systems in mouse models to determine efficacy of the delivery system. The systems also can be tested in human clinical trials.

In instances in which a human HLA class I molecule presents tumor rejection antigens derived from cancer associated nucleic acids, the expression vector may also include a nucleic acid sequence coding for the HLA molecule that presents any particular tumor rejection antigen derived from these nucleic acids and polypeptides. Alternatively, the nucleic acid sequence coding for such a HLA molecule can be contained within a separate expression vector. In a situation where the vector contains both coding sequences, the single vector can be used to transfect a cell which does not normally express either one. Where the coding sequences for a cancer associated antigen precursor and the HLA molecule which presents it are contained on separate expression vectors, the expression vectors can be cotransfected. The cancer associated antigen precursor coding sequence may be used alone, when, e.g. the host cell already expresses a HLA molecule which presents a cancer associated antigen derived from precursor molecules. Of course, there is no limit on the particular host cell which can be used. As the vectors which contain the two coding sequences may be used in any antigen-presenting cells if desired, and the gene for cancer associated antigen precursor can be used in host cells which do not express a HLA molecule which presents a cancer associated antigen. Further, cell-free transcription systems may be used in lieu of cells.

As mentioned above, the invention embraces antisense oligonucleotides that selectively bind to a nucleic acid molecule encoding a cancer associated antigen polypeptide, to reduce the expression of cancer associated antigens. This is desirable in virtually any medical condition wherein a reduction of expression of cancer associated antigens is

desirable, e.g., in the treatment of cancer. This is also useful for *in vitro* or *in vivo* testing of the effects of a reduction of expression of one or more cancer associated antigens.

As used herein, the term "antisense oligonucleotide" or "antisense" describes an oligonucleotide that is an oligoribonucleotide, oligodeoxyribonucleotide, modified oligoribonucleotide, or modified oligodeoxyribonucleotide which hybridizes under physiological conditions to DNA comprising a particular gene or to an mRNA transcript of that gene and, thereby, inhibits the transcription of that gene and/or the translation of that mRNA. The antisense molecules are designed so as to interfere with transcription or translation of a target gene upon hybridization with the target gene or transcript. Those skilled in the art will recognize that the exact length of the antisense oligonucleotide and its degree of complementarity with its target will depend upon the specific target selected, including the sequence of the target and the particular bases which comprise that sequence. It is preferred that the antisense oligonucleotide be constructed and arranged so as to bind selectively with the target under physiological conditions, i.e., to hybridize substantially more to the target sequence than to any other sequence in the target cell under physiological conditions. Based upon the sequences of nucleic acids encoding breast cancer associated antigen, or upon allelic or homologous genomic and/or cDNA sequences, one of skill in the art can easily choose and synthesize any of a number of appropriate antisense molecules for use in accordance with the present invention. In order to be sufficiently selective and potent for inhibition, such antisense oligonucleotides should comprise at least 10 and, more preferably, at least 15 consecutive bases which are complementary to the target, although in certain cases modified oligonucleotides as short as 7 bases in length have been used successfully as antisense oligonucleotides (Wagner et al., *Nature Biotechnol.* 14:840-844, 1996). Most preferably, the antisense oligonucleotides comprise a complementary sequence of 20-30 bases. Although oligonucleotides may be chosen which are antisense to any region of the gene or mRNA transcripts, in preferred embodiments the antisense oligonucleotides correspond to N-terminal or 5' upstream sites such as translation initiation, transcription initiation or promoter sites. In addition, 3'-untranslated regions may be targeted. Targeting to mRNA splicing sites has also been used in the art but may be less preferred if alternative mRNA splicing occurs. In addition, the antisense is targeted, preferably, to sites in which mRNA secondary structure is not expected (see, e.g., Sainio et al., *Cell Mol. Neurobiol.* 14(5):439-457, 1994) and at which proteins are not expected to bind. Finally, although the

listed sequences are cDNA sequences, one of ordinary skill in the art may easily derive the genomic DNA corresponding to the cDNA of a cancer associated antigen. Thus, the present invention also provides for antisense oligonucleotides which are complementary to the genomic DNA corresponding to nucleic acids encoding cancer associated antigens. Similarly, antisense to allelic or homologous cDNAs and genomic DNAs are enabled without undue experimentation.

In one set of embodiments, the antisense oligonucleotides of the invention may be composed of "natural" deoxyribonucleotides, ribonucleotides, or any combination thereof. That is, the 5' end of one native nucleotide and the 3' end of another native nucleotide may be covalently linked, as in natural systems, via a phosphodiester internucleoside linkage. These oligonucleotides may be prepared by art recognized methods which may be carried out manually or by an automated synthesizer. They also may be produced recombinantly by vectors.

In preferred embodiments, however, the antisense oligonucleotides of the invention also may include "modified" oligonucleotides. That is, the oligonucleotides may be modified in a number of ways which do not prevent them from hybridizing to their target but which enhance their stability or targeting or which otherwise enhance their therapeutic effectiveness.

The term "modified oligonucleotide" as used herein describes an oligonucleotide in which (1) at least two of its nucleotides are covalently linked via a synthetic internucleoside linkage (i.e., a linkage other than a phosphodiester linkage between the 5' end of one nucleotide and the 3' end of another nucleotide) and/or (2) a chemical group not normally associated with nucleic acids has been covalently attached to the oligonucleotide. Preferred synthetic internucleoside linkages are phosphorothioates, alkylphosphonates, phosphorodithioates, phosphate esters, alkylphosphonothioates, phosphoramidates, carbamates, carbonates, phosphate triesters, acetamides, carboxymethyl esters and peptides.

The term "modified oligonucleotide" also encompasses oligonucleotides with a covalently modified base and/or sugar. For example, modified oligonucleotides include oligonucleotides having backbone sugars which are covalently attached to low molecular weight organic groups other than a hydroxyl group at the 3' position and other than a phosphate group at the 5' position. Thus modified oligonucleotides may include a 2'-O-alkylated ribose group. In addition, modified oligonucleotides may include sugars such as

arabinose instead of ribose. The present invention, thus, contemplates pharmaceutical preparations containing modified antisense molecules that are complementary to and hybridizable with, under physiological conditions, nucleic acids encoding breast cancer associated antigen polypeptides, together with pharmaceutically acceptable carriers.

Antisense oligonucleotides may be administered as part of a pharmaceutical composition. Such a pharmaceutical composition may include the antisense oligonucleotides in combination with any standard physiologically and/or pharmaceutically acceptable carriers which are known in the art. The compositions should be sterile and contain a therapeutically effective amount of the antisense oligonucleotides in a unit of weight or volume suitable for administration to a patient. The term "pharmaceutically acceptable" means a non-toxic material that does not interfere with the effectiveness of the biological activity of the active ingredients. The term "physiologically acceptable" refers to a non-toxic material that is compatible with a biological system such as a cell, cell culture, tissue, or organism. The characteristics of the carrier will depend on the route of administration. Physiologically and pharmaceutically acceptable carriers include diluents, fillers, salts, buffers, stabilizers, solubilizers, and other materials which are well known in the art, as further described below.

As used herein, a "vector" may be any of a number of nucleic acids into which a desired sequence may be inserted by restriction and ligation for transport between different genetic environments or for expression in a host cell. Vectors are typically composed of DNA although RNA vectors are also available. Vectors include, but are not limited to, plasmids, phagemids and virus genomes. A cloning vector is one which is able to replicate autonomously or integrated in the genome in a host cell, and which is further characterized by one or more endonuclease restriction sites at which the vector may be cut in a determinable fashion and into which a desired DNA sequence may be ligated such that the new recombinant vector retains its ability to replicate in the host cell. In the case of plasmids, replication of the desired sequence may occur many times as the plasmid increases in copy number within the host bacterium or just a single time per host before the host reproduces by mitosis. In the case of phage, replication may occur actively during a lytic phase or passively during a lysogenic phase. An expression vector is one into which a desired DNA sequence may be inserted by restriction and ligation such that it is operably joined to regulatory sequences and may be expressed as an RNA transcript. Vectors may further contain one or more marker sequences suitable for use in the identification of cells which have or have not

been transformed or transfected with the vector. Markers include, for example, genes encoding proteins which increase or decrease either resistance or sensitivity to antibiotics or other compounds, genes which encode enzymes whose activities are detectable by standard assays known in the art (e.g.,  $\beta$ -galactosidase, luciferase or alkaline phosphatase), and genes which visibly affect the phenotype of transformed or transfected cells, hosts, colonies or plaques (e.g., green fluorescent protein). Preferred vectors are those capable of autonomous replication and expression of the structural gene products present in the DNA segments to which they are operably joined.

As used herein, a coding sequence and regulatory sequences are said to be "operably" joined when they are covalently linked in such a way as to place the expression or transcription of the coding sequence under the influence or control of the regulatory sequences. If it is desired that the coding sequences be translated into a functional protein, two DNA sequences are said to be operably joined if induction of a promoter in the 5' regulatory sequences results in the transcription of the coding sequence and if the nature of the linkage between the two DNA sequences does not (1) result in the introduction of a frame-shift mutation, (2) interfere with the ability of the promoter region to direct the transcription of the coding sequences, or (3) interfere with the ability of the corresponding RNA transcript to be translated into a protein. Thus, a promoter region would be operably joined to a coding sequence if the promoter region were capable of effecting transcription of that DNA sequence such that the resulting transcript might be translated into the desired protein or polypeptide.

The precise nature of the regulatory sequences needed for gene expression may vary between species or cell types, but shall in general include, as necessary, 5' non-transcribed and 5' non-translated sequences involved with the initiation of transcription and translation respectively, such as a TATA box, capping sequence, CAAT sequence, and the like. Especially, such 5' non-transcribed regulatory sequences will include a promoter region which includes a promoter sequence for transcriptional control of the operably joined gene. Regulatory sequences may also include enhancer sequences or upstream activator sequences as desired. The vectors of the invention may optionally include 5' leader or signal sequences. The choice and design of an appropriate vector is within the ability and discretion of one of ordinary skill in the art.

Expression vectors containing all the necessary elements for expression are commercially available and known to those skilled in the art. See, e.g., Sambrook et al.,

*Molecular Cloning: A Laboratory Manual*, Second Edition, Cold Spring Harbor Laboratory Press, 1989. Cells are genetically engineered by the introduction into the cells of heterologous DNA (RNA) encoding a breast cancer associated antigen polypeptide or fragment or variant thereof. That heterologous DNA (RNA) is placed under operable control of transcriptional elements to permit the expression of the heterologous DNA in the host cell.

Preferred systems for mRNA expression in mammalian cells are those such as pRc/CMV (available from Invitrogen, Carlsbad, CA) that contain a selectable marker such as a gene that confers G418 resistance (which facilitates the selection of stably transfected cell lines) and the human cytomegalovirus (CMV) enhancer-promoter sequences. Additionally, suitable for expression in primate or canine cell lines is the pCEP4 vector (Invitrogen), which contains an Epstein Barr Virus (EBV) origin of replication, facilitating the maintenance of plasmid as a multicopy extrachromosomal element. Another expression vector is the pEF-BOS plasmid containing the promoter of polypeptide Elongation Factor 1 $\alpha$ , which stimulates efficiently transcription *in vitro*. The plasmid is described by Mishizuma and Nagata (*Nuc. Acids Res.* 18:5322, 1990), and its use in transfection experiments is disclosed by, for example, Demoulin (*Mol. Cell. Biol.* 16:4710-4716, 1996). Still another preferred expression vector is an adenovirus, described by Stratford-Perricaudet, which is defective for E1 and E3 proteins (*J. Clin. Invest.* 90:626-630, 1992). The use of the adenovirus as an Adeno.P1A recombinant for the expression of an antigen is disclosed by Warnier et al., in intradermal injection in mice for immunization against P1A (*Int. J. Cancer*, 67:303-310, 1996). Additional vectors for delivery of nucleic acid are provided below.

The invention also embraces so-called expression kits, which allow the artisan to prepare a desired expression vector or vectors. Such expression kits include at least separate portions of a vector and one or more of the previously discussed cancer associated antigen nucleic acid molecules. Other components may be added, as desired, as long as the previously mentioned nucleic acid molecules, which are required, are included. The invention also includes kits for amplification of a cancer associated antigen nucleic acid, including at least one pair of amplification primers which hybridize to a cancer associated antigen nucleic acid. The primers preferably are 12-32 nucleotides in length and are non-overlapping to prevent formation of "primer-dimers". One of the primers will hybridize to one strand of the cancer associated antigen nucleic acid and the second primer will hybridize to the complementary strand of the cancer associated antigen nucleic acid, in an arrangement which permits



amplification of the cancer associated antigen nucleic acid. Selection of appropriate primer pairs is standard in the art. For example, the selection can be made with assistance of a computer program designed for such a purpose, optionally followed by testing the primers for amplification specificity and efficiency.

The invention also permits the construction of cancer associated antigen gene "knock-outs" in cells and in animals, providing materials for studying certain aspects of cancer and immune system responses to cancer.

The invention also provides isolated polypeptides (including whole proteins and partial proteins) encoded by the foregoing cancer associated antigen nucleic acids. Such polypeptides are useful, for example, alone or as fusion proteins to generate antibodies, as components of an immunoassay or diagnostic assay or as therapeutics. Cancer associated antigen polypeptides can be isolated from biological samples including tissue or cell homogenates, and can also be expressed recombinantly in a variety of prokaryotic and eukaryotic expression systems by constructing an expression vector appropriate to the expression system, introducing the expression vector into the expression system, and isolating the recombinantly expressed protein. Short polypeptides, including antigenic peptides (such as are presented by MHC molecules on the surface of a cell for immune recognition) also can be synthesized chemically using well-established methods of peptide synthesis.

A unique fragment of a cancer associated antigen polypeptide, in general, has the features and characteristics of unique fragments as discussed above in connection with nucleic acids. As will be recognized by those skilled in the art, the size of the unique fragment will depend upon factors such as whether the fragment constitutes a portion of a conserved protein domain. Thus, some regions of cancer associated antigens will require longer segments to be unique while others will require only short segments, typically between 5 and 12 amino acids (e.g. 5, 6, 7, 8, 9, 10, 11 or 12 or more amino acids including each integer up to the full length).

Unique fragments of a polypeptide preferably are those fragments which retain a distinct functional capability of the polypeptide. Functional capabilities which can be retained in a unique fragment of a polypeptide include interaction with antibodies, interaction with other polypeptides or fragments thereof, selective binding of nucleic acids or proteins, and enzymatic activity. One important activity is the ability to act as a signature for identifying the polypeptide. Another is the ability to complex with HLA and to provoke in a human an

immune response. Those skilled in the art are well versed in methods for selecting unique amino acid sequences, typically on the basis of the ability of the unique fragment to selectively distinguish the sequence of interest from non-family members. A comparison of the sequence of the fragment to those on known databases typically is all that is necessary.

The invention embraces variants of the cancer associated antigen polypeptides described above. As used herein, a "variant" of a cancer associated antigen polypeptide is a polypeptide which contains one or more modifications to the primary amino acid sequence of a cancer associated antigen polypeptide. Modifications which create a cancer associated antigen variant can be made to a cancer associated antigen polypeptide 1) to reduce or eliminate an activity of a cancer associated antigen polypeptide; 2) to enhance a property of a cancer associated antigen polypeptide, such as protein stability in an expression system or the stability of protein-protein binding; 3) to provide a novel activity or property to a cancer associated antigen polypeptide, such as addition of an antigenic epitope or addition of a detectable moiety; or 4) to provide equivalent or better binding to an HLA molecule. Modifications to a cancer associated antigen polypeptide are typically made to the nucleic acid which encodes the cancer associated antigen polypeptide, and can include deletions, point mutations, truncations, amino acid substitutions and additions of amino acids or non-amino acid moieties. Alternatively, modifications can be made directly to the polypeptide, such as by cleavage, addition of a linker molecule, addition of a detectable moiety, such as biotin, addition of a fatty acid, and the like. Modifications also embrace fusion proteins comprising all or part of the cancer associated antigen amino acid sequence. One of skill in the art will be familiar with methods for predicting the effect on protein conformation of a change in protein sequence, and can thus "design" a variant cancer associated antigen polypeptide according to known methods. One example of such a method is described by Dahiyat and Mayo in *Science* 278:82-87, 1997, whereby proteins can be designed *de novo*. The method can be applied to a known protein to vary only a portion of the polypeptide sequence. By applying the computational methods of Dahiyat and Mayo, specific variants of a cancer associated antigen polypeptide can be proposed and tested to determine whether the variant retains a desired conformation.

In general, variants include cancer associated antigen polypeptides which are modified specifically to alter a feature of the polypeptide unrelated to its desired physiological activity. For example, cysteine residues can be substituted or deleted to prevent unwanted disulfide

linkages. Similarly, certain amino acids can be changed to enhance expression of a breast cancer associated antigen polypeptide by eliminating proteolysis by proteases in an expression system (e.g., dibasic amino acid residues in yeast expression systems in which KEX2 protease activity is present).

Mutations of a nucleic acid which encode a cancer associated antigen polypeptide preferably preserve the amino acid reading frame of the coding sequence, and preferably do not create regions in the nucleic acid which are likely to hybridize to form secondary structures, such as hairpins or loops, which can be deleterious to expression of the variant polypeptide.

Mutations can be made by selecting an amino acid substitution, or by random mutagenesis of a selected site in a nucleic acid which encodes the polypeptide. Variant polypeptides are then expressed and tested for one or more activities to determine which mutation provides a variant polypeptide with the desired properties. Further mutations can be made to variants (or to non-variant cancer associated antigen polypeptides) which are silent as to the amino acid sequence of the polypeptide, but which provide preferred codons for translation in a particular host. The preferred codons for translation of a nucleic acid in, e.g., *E. coli*, are well known to those of ordinary skill in the art. Still other mutations can be made to the noncoding sequences of a cancer associated antigen gene or cDNA clone to enhance expression of the polypeptide. The activity of variants of cancer associated antigen polypeptides can be tested by cloning the gene encoding the variant cancer associated antigen polypeptide into a bacterial or mammalian expression vector, introducing the vector into an appropriate host cell, expressing the variant cancer associated antigen polypeptide, and testing for a functional capability of the cancer associated antigen polypeptides as disclosed herein. For example, the variant cancer associated antigen polypeptide can be tested for reaction with autologous or allogeneic sera as disclosed in the Examples. Preparation of other variant polypeptides may favor testing of other activities, as will be known to one of ordinary skill in the art.

The skilled artisan will also realize that conservative amino acid substitutions may be made in cancer associated antigen polypeptides to provide functionally equivalent variants of the foregoing polypeptides, i.e., the variants retain the functional capabilities of the cancer associated antigen polypeptides. As used herein, a "conservative amino acid substitution" refers to an amino acid substitution which does not alter the relative charge or size

characteristics of the protein in which the amino acid substitution is made. Variants can be prepared according to methods for altering polypeptide sequence known to one of ordinary skill in the art such as are found in references which compile such methods, e.g. *Molecular Cloning: A Laboratory Manual*, J. Sambrook, et al., eds., Second Edition, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York, 1989, or *Current Protocols in Molecular Biology*, F.M. Ausubel, et al., eds., John Wiley & Sons, Inc., New York. Exemplary functionally equivalent variants of the cancer associated antigen polypeptides include conservative amino acid substitutions of in the amino acid sequences of proteins disclosed herein. Conservative substitutions of amino acids include substitutions made amongst amino acids within the following groups: (a) M, I, L, V; (b) F, Y, W; (c) K, R, H; (d) A, G; (e) S, T; (f) Q, N; and (g) E, D.

For example, upon determining that a peptide derived from a cancer associated antigen polypeptide is presented by an MHC molecule and recognized by CTLs (e.g., as described in the Examples), one can make conservative amino acid substitutions to the amino acid sequence of the peptide, particularly at residues which are thought not to be direct contact points with the MHC molecule. For example, methods for identifying functional variants of HLA class II binding peptides are provided in a published PCT application of Strominger and Wucherpennig (PCT/US96/03182). Peptides bearing one or more amino acid substitutions also can be tested for concordance with known HLA/MHC motifs prior to synthesis using, e.g. the computer program described by D'Amaro and Drijfhout (D'Amaro et al., *Human Immunol.* 43:13-18, 1995; Drijfhout et al., *Human Immunol.* 43:1-12, 1995). The substituted peptides can then be tested for binding to the MHC molecule and recognition by CTLs when bound to MHC. These variants can be tested for improved stability and are useful, *inter alia*, in vaccine compositions.

Conservative amino-acid substitutions in the amino acid sequence of cancer associated antigen polypeptides to produce functionally equivalent variants of cancer associated antigen polypeptides typically are made by alteration of a nucleic acid encoding a cancer associated antigen polypeptide. Such substitutions can be made by a variety of methods known to one of ordinary skill in the art. For example, amino acid substitutions may be made by PCR-directed mutation, site-directed mutagenesis according to the method of Kunkel (Kunkel, *Proc. Nat. Acad. Sci. U.S.A.* 82: 488-492, 1985), or by chemical synthesis of a gene encoding a cancer associated antigen polypeptide. Where amino acid substitutions are made to a small unique

fragment of a cancer associated antigen polypeptide, such as an antigenic epitope recognized by autologous or allogeneic sera or cytolytic T lymphocytes, the substitutions can be made by directly synthesizing the peptide. The activity of functionally equivalent fragments of cancer associated antigen polypeptides can be tested by cloning the gene encoding the altered cancer associated antigen polypeptide into a bacterial or mammalian expression vector, introducing the vector into an appropriate host cell, expressing the altered cancer associated antigen polypeptide, and testing for a functional capability of the cancer associated antigen polypeptides as disclosed herein. Peptides which are chemically synthesized can be tested directly for function, e.g., for binding to antisera recognizing associated antigens.

The invention as described herein has a number of uses, some of which are described elsewhere herein. First, the invention permits isolation of the cancer associated antigen protein molecules. A variety of methodologies well-known to the skilled practitioner can be utilized to obtain isolated cancer associated antigen molecules. The polypeptide may be purified from cells which naturally produce the polypeptide by chromatographic means or immunological recognition. Alternatively, an expression vector may be introduced into cells to cause production of the polypeptide. In another method, mRNA transcripts may be microinjected or otherwise introduced into cells to cause production of the encoded polypeptide. Translation of mRNA in cell-free extracts such as the reticulocyte lysate system also may be used to produce polypeptide. Those skilled in the art also can readily follow known methods for isolating cancer associated antigen polypeptides. These include, but are not limited to, immunochromatography, HPLC, size-exclusion chromatography, ion-exchange chromatography and immune-affinity chromatography.

The isolation and identification of cancer associated antigen genes also makes it possible for the artisan to diagnose a disorder characterized by expression of cancer associated antigens. These methods involve determining expression of one or more cancer associated antigen nucleic acids, and/or encoded cancer associated antigen polypeptides and/or peptides derived therefrom. In the former situation, such determinations can be carried out via any standard nucleic acid determination assay, including the polymerase chain reaction, or assaying with labeled hybridization probes. In the latter situation, such determinations can be carried out by screening patient antisera for recognition of the polypeptide.

The invention also makes it possible isolate proteins which bind to cancer associated antigens as disclosed herein, including antibodies and cellular binding partners of the cancer

associated antigens. Additional uses are described further herein.

The invention also provides, in certain embodiments, "dominant negative" polypeptides derived from cancer associated antigen polypeptides. A dominant negative polypeptide is an inactive variant of a protein, which, by interacting with the cellular machinery, displaces an active protein from its interaction with the cellular machinery or competes with the active protein, thereby reducing the effect of the active protein. For example, a dominant negative receptor which binds a ligand but does not transmit a signal in response to binding of the ligand can reduce the biological effect of expression of the ligand. Likewise, a dominant negative catalytically-inactive kinase which interacts normally with target proteins but does not phosphorylate the target proteins can reduce phosphorylation of the target proteins in response to a cellular signal. Similarly, a dominant negative transcription factor which binds to a promoter site in the control region of a gene but does not increase gene transcription can reduce the effect of a normal transcription factor by occupying promoter binding sites without increasing transcription.

The end result of the expression of a dominant negative polypeptide in a cell is a reduction in function of active proteins. One of ordinary skill in the art can assess the potential for a dominant negative variant of a protein, and using standard mutagenesis techniques to create one or more dominant negative variant polypeptides. For example, given the teachings contained herein of renal cancer associated antigens, especially those which are similar to known proteins which have known activities, one of ordinary skill in the art can modify the sequence of the cancer associated antigens by site-specific mutagenesis, scanning mutagenesis, partial gene deletion or truncation, and the like. See, e.g., U.S. Patent No. 5,580,723 and Sambrook et al., *Molecular Cloning: A Laboratory Manual*, Second Edition, Cold Spring Harbor Laboratory Press, 1989. The skilled artisan then can test the population of mutagenized polypeptides for diminution in a selected and/or for retention of such an activity. Other similar methods for creating and testing dominant negative variants of a protein will be apparent to one of ordinary skill in the art.

The invention also involves agents such as polypeptides which bind to cancer associated antigen polypeptides. Such binding agents can be used, for example, in screening assays to detect the presence or absence of cancer associated antigen polypeptides and complexes of cancer associated antigen polypeptides and their binding partners and in purification protocols to isolated cancer associated antigen polypeptides and complexes of

cancer associated antigen polypeptides and their binding partners. Such agents also can be used to inhibit the native activity of the cancer associated antigen polypeptides, for example, by binding to such polypeptides.

The invention, therefore, embraces peptide binding agents which, for example, can be antibodies or fragments of antibodies having the ability to selectively bind to cancer associated antigen polypeptides. Antibodies include polyclonal and monoclonal antibodies, prepared according to conventional methodology.

Significantly, as is well-known in the art, only a small portion of an antibody molecule, the paratope, is involved in the binding of the antibody to its epitope (see, in general, Clark, W.R. (1986) The Experimental Foundations of Modern Immunology Wiley & Sons, Inc., New York; Roitt, I. (1991) Essential Immunology, 7th Ed., Blackwell Scientific Publications, Oxford). The pFc' and Fc regions, for example, are effectors of the complement cascade but are not involved in antigen binding. An antibody from which the pFc' region has been enzymatically cleaved, or which has been produced without the pFc' region, designated an F(ab')<sub>2</sub> fragment, retains both of the antigen binding sites of an intact antibody. Similarly, an antibody from which the Fc region has been enzymatically cleaved, or which has been produced without the Fc region, designated an Fab fragment, retains one of the antigen binding sites of an intact antibody molecule. Proceeding further, Fab fragments consist of a covalently bound antibody light chain and a portion of the antibody heavy chain denoted Fd. The Fd fragments are the major determinant of antibody specificity (a single Fd fragment may be associated with up to ten different light chains without altering antibody specificity) and Fd fragments retain epitope-binding ability in isolation.

Within the antigen-binding portion of an antibody, as is well-known in the art, there are complementarity determining regions (CDRs), which directly interact with the epitope of the antigen, and framework regions (FRs), which maintain the tertiary structure of the paratope (see, in general, Clark, 1986; Roitt, 1991). In both the heavy chain Fd fragment and the light chain of IgG immunoglobulins, there are four framework regions (FR1 through FR4) separated respectively by three complementarity determining regions (CDR1 through CDR3). The CDRs, and in particular the CDR3 regions, and more particularly the heavy chain CDR3, are largely responsible for antibody specificity.

It is now well-established in the art that the non-CDR regions of a mammalian antibody may be replaced with similar regions of conspecific or heterospecific antibodies

while retaining the epitopic specificity of the original antibody. This is most clearly manifested in the development and use of "humanized" antibodies in which non-human CDRs are covalently joined to human FR and/or Fc/pFc' regions to produce a functional antibody. See, e.g., U.S. patents 4,816,567, 5,225,539, 5,585,089, 5,693,762 and 5,859,205.

Thus, for example, PCT International Publication Number WO 92/04381 teaches the production and use of humanized murine RSV antibodies in which at least a portion of the murine FR regions have been replaced by FR regions of human origin. Such antibodies, including fragments of intact antibodies with antigen-binding ability, are often referred to as "chimeric" antibodies.

Thus, as will be apparent to one of ordinary skill in the art, the present invention also provides for F(ab')<sub>2</sub>, Fab, Fv and Fd fragments; chimeric antibodies in which the Fc and/or FR and/or CDR1 and/or CDR2 and/or light chain CDR3 regions have been replaced by homologous human or non-human sequences; chimeric F(ab')<sub>2</sub> fragment antibodies in which the FR and/or CDR1 and/or CDR2 and/or light chain CDR3 regions have been replaced by homologous human or non-human sequences; chimeric Fab fragment antibodies in which the FR and/or CDR1 and/or CDR2 and/or light chain CDR3 regions have been replaced by homologous human or non-human sequences; and chimeric Fd fragment antibodies in which the FR and/or CDR1 and/or CDR2 regions have been replaced by homologous human or non-human sequences. The present invention also includes so-called single chain antibodies.

Thus, the invention involves polypeptides of numerous size and type that bind specifically to cancer associated antigen polypeptides, and complexes of both cancer associated antigen polypeptides and their binding partners. These polypeptides may be derived also from sources other than antibody technology. For example, such polypeptide binding agents can be provided by degenerate peptide libraries which can be readily prepared in solution, in immobilized form or as phage display libraries. Combinatorial libraries also can be synthesized of peptides containing one or more amino acids. Libraries further can be synthesized of peptoids and non-peptide synthetic moieties.

Phage display can be particularly effective in identifying binding peptides useful according to the invention. Briefly, one prepares a phage library (using e.g. m13, fd, or lambda phage), displaying inserts from 4 to about 80 amino acid residues using conventional procedures. The inserts may represent, for example, a completely degenerate or biased array. One then can select phage-bearing inserts which bind to the cancer associated antigen



polypeptide. This process can be repeated through several cycles of reselection of phage that bind to the cancer associated antigen polypeptide. Repeated rounds lead to enrichment of phage bearing particular sequences. DNA sequence analysis can be conducted to identify the sequences of the expressed polypeptides. The minimal linear portion of the sequence that binds to the cancer associated antigen polypeptide can be determined. One can repeat the procedure using a biased library containing inserts containing part or all of the minimal linear portion plus one or more additional degenerate residues upstream or downstream thereof. Yeast two-hybrid screening methods also may be used to identify polypeptides that bind to the cancer associated antigen polypeptides. Thus, the cancer associated antigen polypeptides of the invention, or a fragment thereof, can be used to screen peptide libraries, including phage display libraries, to identify and select peptide binding partners of the cancer associated antigen polypeptides of the invention. Such molecules can be used, as described, for screening assays, for purification protocols, for interfering directly with the functioning of cancer associated antigen and for other purposes that will be apparent to those of ordinary skill in the art.

As detailed herein, the foregoing antibodies and other binding molecules may be used for example to identify tissues expressing protein or to purify protein. Antibodies also may be coupled to specific diagnostic labeling agents for imaging of cells and tissues that express cancer associated antigens or to therapeutically useful agents according to standard coupling procedures. Diagnostic agents include, but are not limited to, barium sulfate, iocetamic acid, iopanoic acid, ipodate calcium, diatrizoate sodium, diatrizoate meglumine, metrizamide, tyropanoate sodium and radiodiagnostics including positron emitters such as fluorine-18 and carbon-11, gamma emitters such as iodine-123, technitium-99m, iodine-131 and indium-111, nuclides for nuclear magnetic resonance such as fluorine and gadolinium. Other diagnostic agents useful in the invention will be apparent to one of ordinary skill in the art. As used herein, "therapeutically useful agents" include any therapeutic molecule which desirably is targeted selectively to a cell expressing one of the cancer antigens disclosed herein, including antineoplastic agents, radioiodinated compounds, toxins, other cytostatic or cytolytic drugs, and so forth. Antineoplastic therapeutics are well known and include: aminoglutethimide, azathioprine, bleomycin sulfate, busulfan, carmustine, chlorambucil, cisplatin, cyclophosphamide, cyclosporine, cytarabidine, dacarbazine, dactinomycin, daunorubicin, doxorubicin, taxol, etoposide, fluorouracil, interferon- $\alpha$ , lomustine, mercaptopurine,

methotrexate, mitotane, procarbazine HCl, thioguanine, vinblastine sulfate and vincristine sulfate. Additional antineoplastic agents include those disclosed in Chapter 52, Antineoplastic Agents (Paul Calabresi and Bruce A. Chabner), and the introduction thereto, 1202-1263, of Goodman and Gilman's "The Pharmacological Basis of Therapeutics", Eighth Edition, 1990, McGraw-Hill, Inc. (Health Professions Division). Toxins can be proteins such as, for example, pokeweed anti-viral protein, cholera toxin, pertussis toxin, ricin, gelonin, abrin, diphtheria exotoxin, or *Pseudomonas* exotoxin. Toxin moieties can also be high energy-emitting radionuclides such as cobalt-60.

In the foregoing methods, antibodies prepared according to the invention also preferably are specific for the renal cancer associated antigen/MHC complexes described herein.

When "disorder" is used herein, it refers to any pathological condition where the cancer associated antigens are expressed. An example of such a disorder is cancer, breast, colon, gastric, renal, prostate and lung cancers as particular examples.

Samples of tissue and/or cells for use in the various methods described herein can be obtained through standard methods such as tissue biopsy, including punch biopsy and cell scraping, and collection of blood or other bodily fluids by aspiration or other methods.

In certain embodiments of the invention, an immunoreactive cell sample is removed from a subject. By "immunoreactive cell" is meant a cell which can mature into an immune cell (such as a B cell, a helper T cell, or a cytolytic T cell) upon appropriate stimulation. Thus immunoreactive cells include CD34<sup>+</sup> hematopoietic stem cells, immature T cells and immature B cells. When it is desired to produce cytolytic T cells which recognize a cancer associated antigen, the immunoreactive cell is contacted with a cell which expresses a cancer associated antigen under conditions favoring production, differentiation and/or selection of cytolytic T cells; the differentiation of the T cell precursor into a cytolytic T cell upon exposure to antigen is similar to clonal selection of the immune system.

Some therapeutic approaches based upon the disclosure are premised on a response by a subject's immune system, leading to lysis of antigen presenting cells, such as breast cancer cells which present one or more cancer associated antigens. One such approach is the administration of autologous CTLs specific to a cancer associated antigen/MHC complex to a subject with abnormal cells of the phenotype at issue. It is within the ability of one of ordinary skill in the art to develop such CTLs *in vitro*. An example of a method for T cell

differentiation is presented in International Application number PCT/US96/05607. Generally, a sample of cells taken from a subject, such as blood cells, are contacted with a cell presenting the complex and capable of provoking CTLs to proliferate. The target cell can be a transfectant, such as a COS cell. These transfectants present the desired complex of their surface and, when combined with a CTL of interest, stimulate its proliferation. COS cells are widely available, as are other suitable host cells. Specific production of CTL clones is well known in the art. The clonally expanded autologous CTLs then are administered to the subject.

Another method for selecting antigen-specific CTL clones has recently been described (Altman et al., *Science* 274:94-96, 1996; Dunbar et al., *Curr. Biol.* 8:413-416, 1998), in which fluorogenic tetramers of MHC class I molecule/peptide complexes are used to detect specific CTL clones. Briefly, soluble MHC class I molecules are folded *in vitro* in the presence of  $\beta_2$ -microglobulin and a peptide antigen which binds the class I molecule. After purification, the MHC/peptide complex is purified and labeled with biotin. Tetramers are formed by mixing the biotinylated peptide-MHC complex with labeled avidin (e.g. phycoerythrin) at a molar ratio of 4:1. Tetramers are then contacted with a source of CTLs such as peripheral blood or lymph node. The tetramers bind CTLs which recognize the peptide antigen/MHC class I complex. Cells bound by the tetramers can be sorted by fluorescence activated cell sorting to isolate the reactive CTLs. The isolated CTLs then can be expanded *in vitro* for use as described herein.

To detail a therapeutic methodology, referred to as adoptive transfer (Greenberg, *J. Immunol.* 136(5): 1917, 1986; Riddel et al., *Science* 257: 238, 1992; Lynch et al, *Eur. J. Immunol.* 21: 1403-1410, 1991; Kast et al., *Cell* 59: 603-614, 1989), cells presenting the desired complex (e.g., dendritic cells) are combined with CTLs leading to proliferation of the CTLs specific thereto. The proliferated CTLs are then administered to a subject with a cellular abnormality which is characterized by certain of the abnormal cells presenting the particular complex. The CTLs then lyse the abnormal cells, thereby achieving the desired therapeutic goal.

The foregoing therapy assumes that at least some of the subject's abnormal cells present the relevant HLA/cancer associated antigen complex. This can be determined very easily, as the art is very familiar with methods for identifying cells which present a particular HLA molecule, as well as how to identify cells expressing DNA of the pertinent sequences, in

this case a cancer associated antigen sequence. Once cells presenting the relevant complex are identified via the foregoing screening methodology, they can be combined with a sample from a patient, where the sample contains CTLs. If the complex presenting cells are lysed by the mixed CTL sample, then it can be assumed that a cancer associated antigen is being presented, and the subject is an appropriate candidate for the therapeutic approaches set forth *supra*.

Adoptive transfer is not the only form of therapy that is available in accordance with the invention. CTLs can also be provoked *in vivo*, using a number of approaches. One approach is the use of non-proliferative cells expressing the complex. The cells used in this approach may be those that normally express the complex, such as irradiated tumor cells or cells transfected with one or both of the genes necessary for presentation of the complex (i.e. the antigenic peptide and the presenting HLA molecule). Chen et al. (*Proc. Natl. Acad. Sci. USA* 88: 110-114, 1991) exemplifies this approach, showing the use of transfected cells expressing HPVE7 peptides in a therapeutic regime. Various cell types may be used. Similarly, vectors carrying one or both of the genes of interest may be used. Viral or bacterial vectors are especially preferred. For example, nucleic acids which encode a cancer associated antigen polypeptide or peptide may be operably linked to promoter and enhancer sequences which direct expression of the cancer associated antigen polypeptide or peptide in certain tissues or cell types. The nucleic acid may be incorporated into an expression vector. Expression vectors may be unmodified extrachromosomal nucleic acids, plasmids or viral genomes constructed or modified to enable insertion of exogenous nucleic acids, such as those encoding cancer associated antigen, as described elsewhere herein. Nucleic acids encoding a cancer associated antigen also may be inserted into a retroviral genome, thereby facilitating integration of the nucleic acid into the genome of the target tissue or cell type. In these systems, the gene of interest is carried by a microorganism, e.g., a Vaccinia virus, pox virus, herpes simplex virus, retrovirus or adenovirus, and the materials de facto "infect" host cells. The cells which result present the complex of interest, and are recognized by autologous CTLs, which then proliferate.

A similar effect can be achieved by combining the cancer associated antigen or a stimulatory fragment thereof with an adjuvant to facilitate incorporation into antigen presenting cells *in vivo*. The cancer associated antigen polypeptide is processed to yield the peptide partner of the HLA molecule while a cancer associated antigen peptide may be presented without the need for further processing. Generally, subjects can receive an

intradermal injection of an effective amount of the cancer associated antigen. Initial doses can be followed by booster doses, following immunization protocols standard in the art. Preferred cancer associated antigens include those found to react with allogeneic cancer antisera, shown in the examples below.

The invention involves the use of various materials disclosed herein to "immunize" subjects or as "vaccines". As used herein, "immunization" or "vaccination" means increasing or activating an immune response against an antigen. It does not require elimination or eradication of a condition but rather contemplates the clinically favorable enhancement of an immune response toward an antigen. Generally accepted animal models can be used for testing of immunization against cancer using a cancer associated antigen nucleic acid. For example, human cancer cells can be introduced into a mouse to create a tumor, and one or more cancer associated antigen nucleic acids can be delivered by the methods described herein. The effect on the cancer cells (e.g., reduction of tumor size) can be assessed as a measure of the effectiveness of the cancer associated antigen nucleic acid immunization. Of course, testing of the foregoing animal model using more conventional methods for immunization include the administration of one or more cancer associated antigen polypeptides or peptides derived therefrom, optionally combined with one or more adjuvants and/or cytokines to boost the immune response. Methods for immunization, including formulation of a vaccine composition and selection of doses, route of administration and the schedule of administration (e.g. primary and one or more booster doses), are well known in the art. The tests also can be performed in humans, where the end point is to test for the presence of enhanced levels of circulating CTLs against cells bearing the antigen, to test for levels of circulating antibodies against the antigen, to test for the presence of cells expressing the antigen and so forth.

As part of the immunization compositions, one or more cancer associated antigens or stimulatory fragments thereof are administered with one or more adjuvants to induce an immune response or to increase an immune response. An adjuvant is a substance incorporated into or administered with antigen which potentiates the immune response. Adjuvants may enhance the immunological response by providing a reservoir of antigen (extracellularly or within macrophages), activating macrophages and stimulating specific sets of lymphocytes. Adjuvants of many kinds are well known in the art. Specific examples of adjuvants include monophosphoryl lipid A (MPL, SmithKline Beecham), a congener obtained after purification

and acid hydrolysis of *Salmonella minnesota* Re 595 lipopolysaccharide; saponins including QS21 (SmithKline Beecham), a pure QA-21 saponin purified from *Quillja saponaria* extract; DQS21, described in PCT application WO96/33739 (SmithKline Beecham); QS-7, QS-17, QS-18, and QS-L1 (So et al., *Mol. Cells* 7:178-186, 1997); incomplete Freund's adjuvant; complete Freund's adjuvant; montanide; and various water-in-oil emulsions prepared from biodegradable oils such as squalene and/or tocopherol. Preferably, the peptides are administered mixed with a combination of DQS21/MPL. The ratio of DQS21 to MPL typically will be about 1:10 to 10:1, preferably about 1:5 to 5:1 and more preferably about 1:1. Typically for human administration, DQS21 and MPL will be present in a vaccine formulation in the range of about 1 µg to about 100 µg. Other adjuvants are known in the art and can be used in the invention (see, e.g. Goding, *Monoclonal Antibodies: Principles and Practice*, 2nd Ed., 1986). Methods for the preparation of mixtures or emulsions of peptide and adjuvant are well known to those of skill in the art of vaccination.

Other agents which stimulate the immune response of the subject can also be administered to the subject. For example, other cytokines are also useful in vaccination protocols as a result of their lymphocyte regulatory properties. Many other cytokines useful for such purposes will be known to one of ordinary skill in the art, including interleukin-12 (IL-12) which has been shown to enhance the protective effects of vaccines (see, e.g., *Science* 268: 1432-1434, 1995), GM-CSF and IL-18. Thus cytokines can be administered in conjunction with antigens and adjuvants to increase the immune response to the antigens.

There are a number of immune response potentiating compounds that can be used in vaccination protocols. These include costimulatory molecules provided in either protein or nucleic acid form. Such costimulatory molecules include the B7-1 and B7-2 (CD80 and CD86 respectively) molecules which are expressed on dendritic cells (DC) and interact with the CD28 molecule expressed on the T cell. This interaction provides costimulation (signal 2) to an antigen/MHC/TCR stimulated (signal 1) T cell, increasing T cell proliferation and effector function. B7 also interacts with CTLA4 (CD152) on T cells and studies involving CTLA4 and B7 ligands indicate that the B7-CTLA4 interaction can enhance antitumor immunity and CTL proliferation (Zheng P., et al. *Proc. Natl. Acad. Sci. USA* 95 (11):6284-6289 (1998)).

B7 typically is not expressed on tumor cells so they are not efficient antigen presenting cells (APCs) for T cells. Induction of B7 expression would enable the tumor cells to stimulate more efficiently CTL proliferation and effector function. A combination of B7/IL-6/IL-12

costimulation has been shown to induce IFN-gamma and a Th1 cytokine profile in the T cell population leading to further enhanced T cell activity (Gajewski et al., *J. Immunol.*, 154:5637-5648 (1995)). Tumor cell transfection with B7 has been discussed in relation to *in vitro* CTL expansion for adoptive transfer immunotherapy by Wang et al., (*J. Immunol.*, 19:1-8 (1986)). Other delivery mechanisms for the B7 molecule would include nucleic acid (naked DNA) immunization (Kim J., et al. *Nat Biotechnol.*, 15:7:641-646 (1997)) and recombinant viruses such as adeno and pox (Wendtner et al., *Gene Ther.*, 4:7:726-735 (1997)). These systems are all amenable to the construction and use of expression cassettes for the coexpression of B7 with other molecules of choice such as the antigens or fragment(s) of antigens discussed herein (including polytopes) or cytokines. These delivery systems can be used for induction of the appropriate molecules *in vitro* and for *in vivo* vaccination situations. The use of anti-CD28 antibodies to directly stimulate T cells *in vitro* and *in vivo* could also be considered. Similarly, the inducible co-stimulatory molecule ICOS which induces T cell responses to foreign antigen could be modulated, for example, by use of anti-ICOS antibodies (Hutloff et al., *Nature* 397:263-266, 1999).

Lymphocyte function associated antigen-3 (LFA-3) is expressed on APCs and some tumor cells and interacts with CD2 expressed on T cells. This interaction induces T cell IL-2 and IFN-gamma production and can thus complement but not substitute, the B7/CD28 costimulatory interaction (Parra et al., *J. Immunol.*, 158:637-642 (1997), Fenton et al., *J. Immunother.*, 21:2:95-108 (1998)).

Lymphocyte function associated antigen-1 (LFA-1) is expressed on leukocytes and interacts with ICAM-1 expressed on APCs and some tumor cells. This interaction induces T cell IL-2 and IFN-gamma production and can thus complement but not substitute, the B7/CD28 costimulatory interaction (Fenton et al., *J. Immunother.*, 21:2:95-108 (1998)). LFA-1 is thus a further example of a costimulatory molecule that could be provided in a vaccination protocol in the various ways discussed above for B7.

Complete CTL activation and effector function requires Th cell help through the interaction between the Th cell CD40L (CD40 ligand) molecule and the CD40 molecule expressed by DCs (Ridge et al., *Nature*, 393:474 (1998), Bennett et al., *Nature*, 393:478 (1998), Schoenberger et al., *Nature*, 393:480 (1998)). This mechanism of this costimulatory signal is likely to involve upregulation of B7 and associated IL-6/IL-12 production by the DC (APC). The CD40-CD40L interaction thus complements the signal 1 (antigen/MHC-TCR)

and signal 2 (B7-CD28) interactions.

The use of anti-CD40 antibodies to stimulate DC cells directly, would be expected to enhance a response to tumor antigens which are normally encountered outside of a inflammatory context or are presented by non-professional APCs (tumor cells). In these situations Th help and B7 costimulation signals are not provided. This mechanism might be used in the context of antigen pulsed DC based therapies or in situations where Th epitopes have not been defined within known TRA precursors.

A cancer associated antigen polypeptide, or a fragment thereof, also can be used to isolate their native binding partners. Isolation of such binding partners may be performed according to well-known methods. For example, isolated cancer associated antigen polypeptides can be attached to a substrate (e.g., chromatographic media, such as polystyrene beads, or a filter), and then a solution suspected of containing the binding partner may be applied to the substrate. If a binding partner which can interact with cancer associated antigen polypeptides is present in the solution, then it will bind to the substrate-bound cancer associated antigen polypeptide. The binding partner then may be isolated.

It will also be recognized that the invention embraces the use of the cancer associated antigen cDNA sequences in expression vectors, as well as to transfect host cells and cell lines, be these prokaryotic (e.g., *E. coli*), or eukaryotic (e.g., dendritic cells, B cells, CHO cells, COS cells, yeast expression systems and recombinant baculovirus expression in insect cells). Especially useful are mammalian cells such as human, mouse, hamster, pig, goat, primate, etc. They may be of a wide variety of tissue types, and include primary cells and cell lines. Specific examples include keratinocytes, peripheral blood leukocytes, bone marrow stem cells and embryonic stem cells. The expression vectors require that the pertinent sequence, i.e., those nucleic acids described *supra*, be operably linked to a promoter.

The invention also contemplates delivery of nucleic acids, polypeptides or peptides for vaccination. Delivery of polypeptides and peptides can be accomplished according to standard vaccination protocols which are well known in the art. In another embodiment, the delivery of nucleic acid is accomplished by *ex vivo* methods, i.e. by removing a cell from a subject, genetically engineering the cell to include a breast cancer associated antigen, and reintroducing the engineered cell into the subject. One example of such a procedure is outlined in U.S. Patent 5,399,346 and in exhibits submitted in the file history of that patent, all of which are publicly available documents. In general, it involves introduction *in vitro* of a



functional copy of a gene into a cell(s) of a subject, and returning the genetically engineered cell(s) to the subject. The functional copy of the gene is under operable control of regulatory elements which permit expression of the gene in the genetically engineered cell(s). Numerous transfection and transduction techniques as well as appropriate expression vectors are well known to those of ordinary skill in the art, some of which are described in PCT application WO95/00654. *In vivo* nucleic acid delivery using vectors such as viruses and targeted liposomes also is contemplated according to the invention.

In preferred embodiments, a virus vector for delivering a nucleic acid encoding a cancer associated antigen is selected from the group consisting of adenoviruses, adeno-associated viruses, poxviruses including vaccinia viruses and attenuated poxviruses, Semliki Forest virus, Venezuelan equine encephalitis virus, retroviruses, Sindbis virus, and Ty virus-like particle. Examples of viruses and virus-like particles which have been used to deliver exogenous nucleic acids include: replication-defective adenoviruses (e.g., Xiang et al., *Virology* 219:220-227, 1996; Eloit et al., *J. Virol.* 71:5375-5381, 1997; Chengalvala et al., *Vaccine* 15:335-339, 1997), a modified retrovirus (Townsend et al., *J. Virol.* 71:3365-3374, 1997), a nonreplicating retrovirus (Irwin et al., *J. Virol.* 68:5036-5044, 1994), a replication defective Semliki Forest virus (Zhao et al., *Proc. Natl. Acad. Sci. USA* 92:3009-3013, 1995), canarypox virus and highly attenuated vaccinia virus derivative (Paoletti, *Proc. Natl. Acad. Sci. USA* 93:11349-11353, 1996), non-replicative vaccinia virus (Moss, *Proc. Natl. Acad. Sci. USA* 93:11341-11348, 1996), replicative vaccinia virus (Moss, *Dev. Biol. Stand.* 82:55-63, 1994), Venezuelan equine encephalitis virus (Davis et al., *J. Virol.* 70:3781-3787, 1996), Sindbis virus (Pugachev et al., *Virology* 212:587-594, 1995), and Ty virus-like particle (Allsopp et al., *Eur. J. Immunol* 26:1951-1959, 1996). In preferred embodiments, the virus vector is an adenovirus.

Another preferred virus for certain applications is the adeno-associated virus, a double-stranded DNA virus. The adeno-associated virus is capable of infecting a wide range of cell types and species and can be engineered to be replication-deficient. It further has advantages, such as heat and lipid solvent stability, high transduction frequencies in cells of diverse lineages, including hematopoietic cells, and lack of superinfection inhibition thus allowing multiple series of transductions. The adeno-associated virus can integrate into human cellular DNA in a site-specific manner, thereby minimizing the possibility of insertional mutagenesis and variability of inserted gene expression. In addition, wild-type adeno-associated virus

infections have been followed in tissue culture for greater than 100 passages in the absence of selective pressure, implying that the adeno-associated virus genomic integration is a relatively stable event. The adeno-associated virus can also function in an extrachromosomal fashion.

In general, other preferred viral vectors are based on non-cytopathic eukaryotic viruses in which non-essential genes have been replaced with the gene of interest. Non-cytopathic viruses include retroviruses, the life cycle of which involves reverse transcription of genomic viral RNA into DNA with subsequent proviral integration into host cellular DNA.

Adenoviruses and retroviruses have been approved for human gene therapy trials. In general, the retroviruses are replication-deficient (i.e., capable of directing synthesis of the desired proteins, but incapable of manufacturing an infectious particle). Such genetically altered retroviral expression vectors have general utility for the high-efficiency transduction of genes *in vivo*. Standard protocols for producing replication-deficient retroviruses (including the steps of incorporation of exogenous genetic material into a plasmid, transfection of a packaging cell lined with plasmid, production of recombinant retroviruses by the packaging cell line, collection of viral particles from tissue culture media, and infection of the target cells with viral particles) are provided in Kriegler, M., "Gene Transfer and Expression, A Laboratory Manual," W.H. Freeman Co., New York (1990) and Murry, E.J. Ed. "Methods in Molecular Biology," vol. 7, Humana Press, Inc., Clifton, New Jersey (1991).

Preferably the foregoing nucleic acid delivery vectors: (1) contain exogenous genetic material that can be transcribed and translated in a mammalian cell and that can induce an immune response in a host, and (2) contain on a surface a ligand that selectively binds to a receptor on the surface of a target cell, such as a mammalian cell, and thereby gains entry to the target cell.

Various techniques may be employed for introducing nucleic acids of the invention into cells, depending on whether the nucleic acids are introduced *in vitro* or *in vivo* in a host. Such techniques include transfection of nucleic acid- $\text{CaPO}_4$  precipitates, transfection of nucleic acids associated with DEAE, transfection or infection with the foregoing viruses including the nucleic acid of interest, liposome mediated transfection, and the like. For certain uses, it is preferred to target the nucleic acid to particular cells. In such instances, a vehicle used for delivering a nucleic acid of the invention into a cell (e.g., a retrovirus, or other virus; a liposome) can have a targeting molecule attached thereto. For example, a molecule such as an antibody specific for a surface membrane protein on the target cell or a

ligand for a receptor on the target cell can be bound to or incorporated within the nucleic acid delivery vehicle. Preferred antibodies include antibodies which selectively bind a cancer associated antigen, alone or as a complex with a MHC molecule. Especially preferred are monoclonal antibodies. Where liposomes are employed to deliver the nucleic acids of the invention, proteins which bind to a surface membrane protein associated with endocytosis may be incorporated into the liposome formulation for targeting and/or to facilitate uptake. Such proteins include capsid proteins or fragments thereof tropic for a particular cell type, antibodies for proteins which undergo internalization in cycling, proteins that target intracellular localization and enhance intracellular half life, and the like. Polymeric delivery systems also have been used successfully to deliver nucleic acids into cells, as is known by those skilled in the art. Such systems even permit oral delivery of nucleic acids.

When administered, the therapeutic compositions of the present invention can be administered in pharmaceutically acceptable preparations. Such preparations may routinely contain pharmaceutically acceptable concentrations of salt, buffering agents, preservatives, compatible carriers, supplementary immune potentiating agents such as adjuvants and cytokines and optionally other therapeutic agents.

The therapeutics of the invention can be administered by any conventional route, including injection or by gradual infusion over time. The administration may, for example, be oral, intravenous, intraperitoneal, intramuscular, intracavity, subcutaneous, or transdermal. When antibodies are used therapeutically, a preferred route of administration is by pulmonary aerosol. Techniques for preparing aerosol delivery systems containing antibodies are well known to those of skill in the art. Generally, such systems should utilize components which will not significantly impair the biological properties of the antibodies, such as the paratope binding capacity (see, for example, Sciarra and Cutie, "Aerosols," in Remington's Pharmaceutical Sciences, 18th edition, 1990, pp 1694-1712; incorporated by reference). Those of skill in the art can readily determine the various parameters and conditions for producing antibody aerosols without resort to undue experimentation. When using antisense preparations of the invention, slow intravenous administration is preferred.

The compositions of the invention are administered in effective amounts. An "effective amount" is that amount of a cancer associated antigen composition that alone, or together with further doses, produces the desired response, e.g. increases an immune response to the cancer associated antigen. In the case of treating a particular disease or condition

characterized by expression of one or more cancer associated antigens, such as renal cancer, the desired response is inhibiting the progression of the disease. This may involve only slowing the progression of the disease temporarily, although more preferably, it involves halting the progression of the disease permanently. This can be monitored by routine methods or can be monitored according to diagnostic methods of the invention discussed herein. The desired response to treatment of the disease or condition also can be delaying the onset or even preventing the onset of the disease or condition.

Such amounts will depend, of course, on the particular condition being treated, the severity of the condition, the individual patient parameters including age, physical condition, size and weight, the duration of the treatment, the nature of concurrent therapy (if any), the specific route of administration and like factors within the knowledge and expertise of the health practitioner. These factors are well known to those of ordinary skill in the art and can be addressed with no more than routine experimentation. It is generally preferred that a maximum dose of the individual components or combinations thereof be used, that is, the highest safe dose according to sound medical judgment. It will be understood by those of ordinary skill in the art, however, that a patient may insist upon a lower dose or tolerable dose for medical reasons, psychological reasons or for virtually any other reasons.

The pharmaceutical compositions used in the foregoing methods preferably are sterile and contain an effective amount of cancer associated antigen or nucleic acid encoding cancer associated antigen for producing the desired response in a unit of weight or volume suitable for administration to a patient. The response can, for example, be measured by determining the immune response following administration of the cancer associated antigen composition via a reporter system by measuring downstream effects such as gene expression, or by measuring the physiological effects of the cancer associated antigen composition, such as regression of a tumor or decrease of disease symptoms. Other assays will be known to one of ordinary skill in the art and can be employed for measuring the level of the response.

The doses of cancer associated antigen compositions (e.g., polypeptide, peptide, antibody, cell or nucleic acid) administered to a subject can be chosen in accordance with different parameters, in particular in accordance with the mode of administration used and the state of the subject. Other factors include the desired period of treatment. In the event that a response in a subject is insufficient at the initial doses applied, higher doses (or effectively higher doses by a different, more localized delivery route) may be employed to the extent that

patient tolerance permits.

In general, for treatments for eliciting or increasing an immune response, doses of cancer associated antigen are formulated and administered in doses between 1 ng and 1 mg, and preferably between 10 ng and 100 µg, according to any standard procedure in the art. Where nucleic acids encoding cancer associated antigen of variants thereof are employed, doses of between 1 ng and 0.1 mg generally will be formulated and administered according to standard procedures. Other protocols for the administration of cancer associated antigen compositions will be known to one of ordinary skill in the art, in which the dose amount, schedule of injections, sites of injections, mode of administration (e.g., intra-tumoral) and the like vary from the foregoing. Administration of cancer associated antigen compositions to mammals other than humans, e.g. for testing purposes or veterinary therapeutic purposes, is carried out under substantially the same conditions as described above.

Where cancer associated antigen peptides are used for vaccination, modes of administration which effectively deliver the cancer associated antigen and adjuvant, such that an immune response to the antigen is increased, can be used. For administration of a cancer associated antigen peptide in adjuvant, preferred methods include intradermal, intravenous, intramuscular and subcutaneous administration. Although these are preferred embodiments, the invention is not limited by the particular modes of administration disclosed herein. Standard references in the art (e.g., *Remington's Pharmaceutical Sciences*, 18th edition, 1990) provide modes of administration and formulations for delivery of immunogens with adjuvant or in a non-adjuvant carrier.

When administered, the pharmaceutical preparations of the invention are applied in pharmaceutically-acceptable amounts and in pharmaceutically-acceptable compositions. The term "pharmaceutically acceptable" means a non-toxic material that does not interfere with the effectiveness of the biological activity of the active ingredients. Such preparations may routinely contain salts, buffering agents, preservatives, compatible carriers, and optionally other therapeutic agents. When used in medicine, the salts should be pharmaceutically acceptable, but non-pharmaceutically acceptable salts may conveniently be used to prepare pharmaceutically-acceptable salts thereof and are not excluded from the scope of the invention. Such pharmacologically and pharmaceutically-acceptable salts include, but are not limited to, those prepared from the following acids: hydrochloric, hydrobromic, sulfuric, nitric, phosphoric, maleic, acetic, salicylic, citric, formic, malonic, succinic, and the like.

Also, pharmaceutically-acceptable salts can be prepared as alkaline metal or alkaline earth salts, such as sodium, potassium or calcium salts.

A renal cancer associated antigen composition may be combined, if desired, with a pharmaceutically-acceptable carrier. The term "pharmaceutically-acceptable carrier" as used herein means one or more compatible solid or liquid fillers, diluents or encapsulating substances which are suitable for administration into a human. The term "carrier" denotes an organic or inorganic ingredient, natural or synthetic, with which the active ingredient is combined to facilitate the application. The components of the pharmaceutical compositions also are capable of being co-mingled with the molecules of the present invention, and with each other, in a manner such that there is no interaction which would substantially impair the desired pharmaceutical efficacy.

The pharmaceutical compositions may contain suitable buffering agents, including: acetic acid in a salt; citric acid in a salt; boric acid in a salt; and phosphoric acid in a salt.

The pharmaceutical compositions also may contain, optionally, suitable preservatives, such as: benzalkonium chloride; chlorobutanol; parabens and thimerosal.

The pharmaceutical compositions may conveniently be presented in unit dosage form and may be prepared by any of the methods well-known in the art of pharmacy. All methods include the step of bringing the active agent into association with a carrier which constitutes one or more accessory ingredients. In general, the compositions are prepared by uniformly and intimately bringing the active compound into association with a liquid carrier, a finely divided solid carrier, or both, and then, if necessary, shaping the product.

Compositions suitable for oral administration may be presented as discrete units, such as capsules, tablets, lozenges, each containing a predetermined amount of the active compound. Other compositions include suspensions in aqueous liquids or non-aqueous liquids such as a syrup, elixir or an emulsion.

Compositions suitable for parenteral administration conveniently comprise a sterile aqueous or non-aqueous preparation of breast cancer associated antigen polypeptides or nucleic acids, which is preferably isotonic with the blood of the recipient. This preparation may be formulated according to known methods using suitable dispersing or wetting agents and suspending agents. The sterile injectable preparation also may be a sterile injectable solution or suspension in a non-toxic parenterally-acceptable diluent or solvent, for example, as a solution in 1,3-butane diol. Among the acceptable vehicles and solvents that may be

employed are water, Ringer's solution, and isotonic sodium chloride solution. In addition, sterile, fixed oils are conventionally employed as a solvent or suspending medium. For this purpose any bland fixed oil may be employed including synthetic mono-or di-glycerides. In addition, fatty acids such as oleic acid may be used in the preparation of injectables. Carrier formulation suitable for oral, subcutaneous, intravenous, intramuscular, etc. administrations can be found in *Remington's Pharmaceutical Sciences*, Mack Publishing Co., Easton, PA.

As used herein with respect to nucleic acids, the term "isolated" means: (i) amplified *in vitro* by, for example, polymerase chain reaction (PCR); (ii) recombinantly produced by cloning; (iii) purified, as by cleavage and gel separation; or (iv) synthesized by, for example, chemical synthesis. An isolated nucleic acid is one which is readily manipulable by recombinant DNA techniques well known in the art. Thus, a nucleotide sequence contained in a vector in which 5' and 3' restriction sites are known or for which polymerase chain reaction (PCR) primer sequences have been disclosed is considered isolated but a nucleic acid sequence existing in its native state in its natural host is not. An isolated nucleic acid may be substantially purified, but need not be. For example, a nucleic acid that is isolated within a cloning or expression vector is not pure in that it may comprise only a tiny percentage of the material in the cell in which it resides. Such a nucleic acid is isolated, however, as the term is used herein because it is readily manipulable by standard techniques known to those of ordinary skill in the art. An isolated nucleic acid as used herein is not a naturally occurring chromosome.

As used herein with respect to polypeptides, "isolated" means separated from its native environment and present in sufficient quantity to permit its identification or use. Isolated, when referring to a protein or polypeptide, means, for example: (i) selectively produced by expression cloning or (ii) purified as by chromatography or electrophoresis. Isolated proteins or polypeptides may, but need not be, substantially pure. The term "substantially pure" means that the proteins or polypeptides are essentially free of other substances with which they may be found in nature or *in vivo* systems to an extent practical and appropriate for their intended use. Substantially pure polypeptides may be produced by techniques well known in the art. Because an isolated protein may be admixed with a pharmaceutically acceptable carrier in a pharmaceutical preparation, the protein may comprise only a small percentage by weight of the preparation. The protein is nonetheless isolated in that it has been separated from the substances with which it may be associated in living systems, i.e. isolated from other proteins.

### Examples

#### Example 1: SEREX screening of renal cancer cell line 1973/10.4

A standard cDNA library was prepared using 5 µg of poly A<sup>+</sup> RNA derived from the cell line 1973/10.4. A primary (unamplified) cDNA library was immunoscreened (5 x 10<sup>5</sup> clones per library) by standard SEREX methodology, with absorbed autologous patient serum at 1:200 dilution [Sahin, U. et al., *Proc Natl Acad Sci USA* 92:11810-3 (1995); Chen, Y.T. et al. *Proc Natl Acad Sci USA*. 94:1914-8 (1997)]. Excluding false-positive clones encoding immunoglobulin gene fragments, clones were purified and sequence analyzed. Comparisons of the sequences showed that these clones represented cDNAs from 22 distinct genes, designated NY-REN-45 through NY-LU-66 (Table A and Sequence Listing (SEQ ID NOs:1-21)). A homology search through the GenBank/EMBO databases revealed that 14 of the 22 genes corresponded to previously known molecules, and 8 others were unknown genes, with sequence identity limited only to short segments of known genes or to expressed sequence tags (ESTs).

Analysis of isolated clones:

#### I. NY-REN clones which are known gene products

Designation	Gene/Sequence Identity	Accession Number	SEQ ID NO
NY-REN-46	lactate dehydrogenase B	Y00711	22
NY-REN-47	ERK tyrosine kinase	D31661	23
NY-REN-48	PINCH protein	U09284	24
NY-REN-51	BBP/53BP2	U58334	25
NY-REN-52	steroid receptor coactivator	U59302	26
NY-REN-53	KIAA0336 mRNA tag	AB002334	27
NY-REN-54	E6 oncogenic protein-associated protein	X98033	28
NY-REN-55	murine NEK1 protein kinase homologue	S45828	29



NY-REN-56	6-phospho-fructokinase	D49817	30
NY-REN-59	lactate dehydrogenase A	X02152	31
NY-REN-61	KIAA0081 mRNA tag	D42039	32
NY-REN-63	DDB p127-associated protein	AF035950	33
NY-REN-65	HREV107 protein	X92814	34
NY-REN-66	acidic ribosomal phosphoprotein 2	M17887	35

## II. Novel gene products

<u>Clone</u>	<u>SEQ ID NO:</u>	<u>Size</u>	<u>Tissue mRNA expression</u>	<u>Protein SEQ ID NO:</u>
NY-REN-45	1	4.0kb	Ubiquitous	12
NY-REN-49	2	1.1kb	Ubiquitous	13
NY-REN-50	3	1.8kb	Ubiquitous	14
NY-REN-57	4,5	2.9kb	Ubiquitous	15, 16
NY-REN-58	6	1.9kb	Ubiquitous	17
NY-REN-60	7	4.0kb	Ubiquitous	18
NY-REN-62	8,9	2.7kb	Ubiquitous	19, 20
NY-REN-64	10,11	3.0kb	Ubiquitous	21

## III. Clones which react with autologous sera only:

NY-REN-47

NY-REN-49

NY-REN-50

## IV. Clones which react with sera from normal control donors

Frequency of sera reactivity

<u>Clone</u>	<u>normal</u>	<u>cancer patient</u>
NY-REN-46	4/4	6/14
NY-REN-48	14/14	17/17
NY-REN-51	1/12	3/17

- 52 -

NY-REN-52	4/12	7/17
NY-REN-53	4/19	10/31
NY-REN-54	5/8	7/7
NY-REN-55	3/19	6/31
NY-REN-56	3/19	7/31
NY-REN-57	1/19	3/31
NY-REN-58	1/12	2/17
NY-REN-59	1/19	4/31
NY-REN-61	3/19	5/31
NY-REN-62	1/19	4/31
NY-REN-63	2/19	12/31
NY-REN-64	3/19	6/31
NY-REN-65	2/19	2/31

V. Clones which react with sera from cancer patients only (failed to react with 19 normal patient serum samples). These clones are preferred for therapeutic and diagnostic applications.

	<u>Frequency of reactivity</u>
NY-REN-32	3/31
NY-REN-45	3/31
NY-REN-57	2/31
NY-REN-60	5/31
NY-REN-66	2/31

#### VI. Additional allogeneic screening of NY-REN renal SEREX clones

Renal SEREX clones were tested for reactivity with sera from the normal and various cancer patients listed below.

---

Sera

<u>Clone</u>	<u>normal</u>	<u>colon</u>	<u>renal</u>	<u>lung</u>	<u>breast</u>
NY-REN-3	0/26	7/37	8/32	0/23	1/26
NY-REN-12	0/19	0/16	3/32	0/15	0/16
NY-REN-19	0/19	0/16	2/32	0/15	0/16
NY-REN-21	0/16	3/16	3/32	1/15	0/16
NY-REN-25	0/15	0/16	5/32	0/15	0/16
NY-REN-31	0/14	0/16	5/32	0/15	0/16
NY-REN-32	0/14	2/16	3/32	0/15	0/16
NY-REN-37	0/15	0/16	2/32	0/15	0/16
NY-REN-45	0/14	0/16	3/32	1/15	0/16
NY-REN-57	0/19	0/16	2/32	0/15	0/16
NY-REN-60	0/19	0/16	7/32	0/15	0/16
NY-REN-66	0/19	0/16	2/32	0/15	0/16

**Example 2: SEREX screening of tumor cells and testis libraries**

Standard cDNA libraries were prepared using poly A<sup>+</sup> RNA derived from the various cancer sources, as well as testis. The cDNA libraries were immunoscreened by standard SEREX methodology, with absorbed autologous patient serum [Sahin, U. et al., *Proc Natl Acad Sci USA* 92:11810-3 (1995); Chen, Y.T. et al. *Proc Natl Acad Sci USA*. 94:1914-8 (1997)]. Excluding false-positive clones encoding immunoglobulin gene fragments, clones were purified and sequence analyzed. Comparisons of the sequences showed that these clones all were substantially identical to the kinectin cDNA (GenBank Accession number L25616).

**Table A: Kinectin clones isolated by SEREX**

<u>Clone designation</u>	<u>SEQ ID NO</u>	<u>Library</u>	<u>Patient sera for screen</u>
NGO-St-47	36	Gastric tumor	autologous
TS-64-5'	37	Testis	seminoma
HOM-TSOv3-41 (5')	38	Testis	seminoma
HOM-TSOv3-41 (3')	39	Testis	seminoma
HOM-HD2-2(3')	40	Hodgkin's disease	autologous
HOM-HD2-232(3')	41	Hodgkin's disease	autologous
Thy4 (3')	42	Thyroid tumor	autologous

Thy4 (3')	42	Thyroid tumor	autologous
Thy5 (3')	43	Thyroid tumor	autologous
Thy8 (3')	44	Thyroid tumor	autologous
Thy10 (3')	45	Thyroid tumor	autologous
NGO-Br-1	46	Breast cancer	autologous

In addition, a hepatocarcinoma library screened with autologous sera also identified several clones which were sequenced and found to be substantially identical to kinectin.

#### Allotyping (gastric cancer)

12/12 gastric cancer patient sera recognized kinectin

0/27 normal individual sera recognized kinectin

Therefore it was determined that recognition of kinectin was diagnostic for cancer patients.

#### **Example 3: Preparation of recombinant cancer associated antigens**

To facilitate screening of patients' sera for antibodies reactive with cancer associated antigens, for example by ELISA, recombinant proteins are prepared according to standard procedures. In one method, the clones encoding cancer associated antigens are subcloned into a baculovirus expression vector, and the recombinant expression vectors are introduced into appropriate insect cells. Baculovirus/insect cloning systems are preferred because post-translational modifications are carried out in the insect cells. Another preferred eukaryotic system is the *Drosophila* Expression System from Invitrogen. Clones which express high amounts of the recombinant protein are selected and used to produce the recombinant proteins. The recombinant proteins are tested for antibody recognition using serum from the patient which was used to isolate the particular clone, or in the case of cancer associated antigens recognized by allogeneic sera, by the sera from any of the patients used to isolate the clones or sera which recognize the clones' gene products.

Alternatively, the cancer associated antigen clones are inserted into a prokaryotic expression vector for production of recombinant proteins in bacteria. Other systems, including yeast expression systems and mammalian cell culture systems also can be used.

#### **Example 4: Preparation of antibodies to cancer associated antigens**

The recombinant cancer associated antigens produced as in Example 4 above are used to generate polyclonal antisera and monoclonal antibodies according to standard procedures. The antisera and antibodies so produced are tested for correct recognition of the cancer associated antigens by using the antisera/antibodies in assays of cell extracts of patients known to express the particular cancer associated antigen (e.g. an ELISA assay). These antibodies can be used for experimental purposes (e.g. localization of the cancer associated antigens, immunoprecipitations, Western blots, etc.) as well as diagnostic purposes (e.g., testing extracts of tissue biopsies, testing for the presence of cancer associated antigens).

**Example 5: Expression of cancer associated antigens in cancers of similar and different origin.**

The expression of one or more of the cancer associated antigens is tested in a range of tumor samples to determine which, if any, other malignancies should be diagnosed and/or treated by the methods described herein. Tumor cell lines and tumor samples are tested for cancer associated antigen expression, preferably by RT-PCR according to standard procedures. Northern blots also are used to test the expression of the cancer associated antigens. Antibody based assays, such as ELISA and western blot, also can be used to determine protein expression. A preferred method of testing expression of cancer associated antigens (in other cancers and in additional same type cancer patients) is allogeneic serotyping using a modified SEREX protocol (as described above).

In all of the foregoing, extracts from the tumors of patients who provided sera for the initial isolation of the cancer associated antigens are used as positive controls. The cells containing recombinant expression vectors described in the Examples above also can be used as positive controls.

The results generated from the foregoing experiments provide panels of multiple cancer associated nucleic acids and/or polypeptides for use in diagnostic (e.g. determining the existence of cancer, determining the prognosis of a patient undergoing therapy, etc.) and therapeutic methods (e.g., vaccine composition, etc.).

**Example 6: HLA typing of patients positive for cancer associated antigens**

To determine which HLA molecules present peptides derived from the cancer associated antigens of the invention, cells of the patients which express the cancer associated

antigens are HLA typed. Peripheral blood lymphocytes are taken from the patient and typed for HLA class I or class II, as well as for the particular subtype of class I or class II. Tumor biopsy samples also can be used for typing. HLA typing can be carried out by any of the standard methods in the art of clinical immunology, such as by recognition by specific monoclonal antibodies, or by HLA allele-specific PCR (e.g. as described in WO97/31126).

**Example 7: Characterization of cancer associated antigen peptides presented by MHC class I and class II molecules.**

Antigens which provoke an antibody response in a subject may also provoke a cell-mediated immune response. Cells process proteins into peptides for presentation on MHC class I or class II molecules on the cell surface for immune surveillance. Peptides presented by certain MHC/HLA molecules generally conform to motifs. These motifs are known in some cases, and can be used to screen the renal cancer associated antigens for the presence of potential class I and/or class II peptides. Summaries of class I and class II motifs have been published (e.g., Rammensee et al., *Immunogenetics* 41:178-228, 1995). Based on the results of experiments such as those described above, the HLA types which present the individual breast cancer associated antigens are known. Motifs of peptides presented by these HLA molecules thus are preferentially searched.

One also can search for class I and class II motifs using computer algorithms. For example, computer programs for predicting potential CTL epitopes based on known class I motifs has been described (*see, e.g.*, Parker et al., *J. Immunol.* 152:163, 1994; D'Amato et al., *Human Immunol.* 43:13-18, 1995; Drijfhout et al., *Human Immunol.* 43:1-12, 1995). Computer programs for predicting potential T cell epitopes based on known class II motifs has also been described (*see, e.g.* Sturniolo et al., *Nat Biotechnol* 17(6):555-61, 1999). HLA binding predictions can conveniently be made using an algorithm available via the Internet on the National Institutes of Health World Wide Web site at URL <http://bimas.dcrt.nih.gov>. See also the website of: SYFPEITHI: An Internet Database for MHC Ligands and Peptide Motifs (access via <http://www.uni-tuebingen.de/uni/kxi/> or <http://134.2.96.221/scripts/hlaserver.dll/EpPredict.htm>). Methods for determining HLA class II peptides and making substitutions thereto are also known (e.g. Strominger and Wucherpfennig (PCT/US96/03182)).

**Example 8: Identification of the portion of a cancer associated polypeptide encoding an antigen**

To determine if the cancer associated antigens isolated as described above can provoke a cytolytic T lymphocyte response, the following method is performed. CTL clones are generated by stimulating the peripheral blood lymphocytes (PBLs) of a patient with autologous normal cells transfected with one of the clones encoding a cancer associated antigen polypeptide or with irradiated PBLs loaded with synthetic peptides corresponding to the putative protein and matching the consensus for the appropriate HLA class I molecule (as described above) to localize an antigenic peptide within the cancer associated antigen clone (see, e.g., Knuth et al., *Proc. Natl. Acad. Sci. USA* 81:3511-3515, 1984; van der Bruggen et al., *Eur. J. Immunol.* 24:3038-3043, 1994). These CTL clones are screened for specificity against COS cells transfected with the cancer associated antigen clone and autologous HLA alleles as described by Brichard et al. (*Eur. J. Immunol.* 26:224-230, 1996). CTL recognition of a cancer associated antigen is determined by measuring release of TNF from the cytolytic T lymphocyte or by  $^{51}\text{Cr}$  release assay (Herin et al., *Int. J. Cancer* 39:390-396, 1987). If a CTL clone specifically recognizes a transfected COS cell, then shorter fragments of the cancer associated antigen clone transfected in that COS cell are tested to identify the region of the gene that encodes the peptide. Fragments of the cancer associated antigen clone are prepared by exonuclease III digestion or other standard molecular biology methods. Synthetic peptides are prepared to confirm the exact sequence of the antigen.

Optionally, shorter fragments of cancer associated antigen cDNAs are generated by PCR. Shorter fragments are used to provoke TNF release or  $^{51}\text{Cr}$  release as above.

Synthetic peptides corresponding to portions of the shortest fragment of the cancer associated antigen clone which provokes TNF release are prepared. Progressively shorter peptides are synthesized to determine the optimal cancer associated antigen tumor rejection antigen peptides for a given HLA molecule.

A similar method is performed to determine if the cancer associated antigen contains one or more HLA class II peptides recognized by T cells. One can search the sequence of the cancer associated antigen polypeptides for HLA class II motifs as described above. In contrast to class I peptides, class II peptides are presented by a limited number of cell types. Thus for these experiments, dendritic cells or B cell clones which express HLA class II molecules preferably are used.

**Table 1: Sequence homologies**

SEQ ID NO.: 1  
 AB002794, U46118RNU46118U19482AF026216AB002739, AB002730, AB002728,  
 AF058796, AB002777, AF020187, AF009411, AB015609, AF006627, M95076, U43527,  
 AB002741, U25846, AF006628, AF019043, D83352, U10355, M99575, U48288,  
 AF090440, AB007504, U82480, Y15794, AJ005572, AF029349, L10111, S80963,  
 U38894, L41731, AF022733, AJ225108, Y11879, AF001688, U33214, Z97178,  
 AF009413, AF019907, AF016371, X71980, AF001522, M77169, AF023132, D83476,  
 AJ009675, U90554, AF069324, AF048691, Z34799, AF004947, U60149, AF022732,  
 AF019887, L02937, U55848, AF011331, AF045770, AF043533, X69524, M32882,  
 U08214, U50847, AF017364, U35364, Y14339, AJ005969, U90555, U44430, U74296,  
 AF001501, AJ223316, D78609, U41060, Z54362, AF034387, U62398, AB005545,  
 AB002533, Y13865, AF009959, L02938, U10555, AL010246, AC003041, U37699,  
 S75970, AF071010, Z81311, AB003681, X16353, AB002731, X53081, D63450,  
 D10911, D90115, X97065, Z82275, AB016891, Y09455, X77990, , W58357, W07820,  
 AA188593, W81046, AA858164, AI018124, AI139112, AA769634, W00437, D80849,  
 AA448160, N98650, T31293, H15307, H51146, N29314, AA770301, AA187822,  
 AA978299, T31823, AA936410, AA993194, N41386, W45601, W81099, C05691,  
 D60153, AA780119, AA929004, R68608, Z43271, N37024, AA974718, AA928663,  
 T30120, AA936583, H51109, N20251, W92917, T36035, AA480197, Z24908, N79450,  
 AA872019, AA902275, N59810, R13443, AA740162, AA937759, N72168, AA970708,  
 T36257, AA433833, H15700, D51185, AA249138, R37356, W93028, Z28641,  
 AA371494, N72757, AA719126, N56164, T31790, D57384, Z41959, N58984,  
 AA847848, AA587009, Z39343, AA654834, AA505490, Z38243, D20344, AA252395,  
 AA025593, H70133, R09101, AA159862, AI032981, Z28355, C75020, AI139642,  
 D82421, AI126922, C75170, AI016032, C18748, C75478, AI129334, C75472,  
 AA310765, D63057, C05952, AA357303, C16591, D82132, L48852, D82799, C75176,  
 AI124552, AA669404, U30155, C75118, AA374918, C75108, C05853, U30151,  
 D57346, AI127548, AA918527, AA317816, AA573490, R21699, AA917928, R36311,  
 AA361522, AA701252, AI085492, H44387, AA156256, AA587935, AA976510,  
 AA515269, W73374, T27986, N34493, AA737770, N32609, N32612, H64420, ,  
 AA415243, AA413717, AA117350, AA242502, AA117343, AA545256, AA795651,  
 AA106372, U31322, AA681967, AA221922, AA600546, AA050610, AU018628, C80932,  
 AA920654, AA863834, AI099036, AA183239, AI115182, AA590910, W65628,  
 AA162291, , AA109440, AI052952, AA999324, AI105714, AI026280, AI072678,  
 AA964820, AA754198, AA944557, AI045710, C94989, AA471630, AA933231,  
 AA509077, AA257557, AA509328, AA109365, AA963207, AA435473, AA999306,  
 AA406684, AI105662, AA509174, AI108597, AA109374, AA925182, AA471671,  
 AA088161, AA752812, AA626989, AA406924, AA840999, AA509033, C94899,  
 AA406875, AA842672, AA123619, AA072471, AA509339, AA933108, AI110259,  
 AA842135, N99262, AA661430, AA842512, AA180648, AI066048, AA559515,  
 AA406761, AA109423, AF074736, AA257676, D87312, AA257488, AA114390,  
 AI087701, W06665, W32852, AA123649, AA842505, AA509330, AA908025, AA257279,  
 AI105717, AA406673, AA627018, AA406850, AA161668, AA114453, AA109372,  
 AA257572, AA803997, AA406943, AA257495, AA842019, AA661368, AI082949,  
 AA842888, W06539, AA118223, AA406986, AA471717, AI096212, AA257707,  
 AA257649, AI053126, AA508950, AA713366, W69049, AA965381, AA208680,  
 AI057997, AA161707, AI058071, AA675858, AA257513, AI083337, AA406908,  
 AI083256, AA659956, AA842532, AA754036, AA753165, W18199, AI109620,  
 AI058022, AI064026, AA471431, AA032116, AA754544, AI108146, C42045,  
 AA843025, AI109103, AA509237, AA161620, AA257665, AA471605, W84932,  
 AA841358, AA843040, AA471695, AA003471, AA840977, AA509267, AA841711,  
 AA471488, AA257699, N98079, AA454318, W04102, AA842874, N99754, AA509307,  
 AA751845, AA661358, AA508954, AA406746, AA842720, W23363, AA508951,  
 AI068913, AA417420, AI108220, AA508946, AA471497, AA842911, AA842501,  
 AA123614, AA509218, AA547812, AA752986, AA752003, AA406839, AA508993,  
 AA161747, AA509008, AA180623, AA508986, AA842627, AA753129, AA509214,  
 AA406733, AI105737, N74818, AA753300, AA430818, AA417415, AA114426, R46936,  
 AA979824, AA246112, AA753138, AI058077, AA842464, AA966639, AA842265,  
 AA509025, AA601853, W15128, AA660039, AA842275, AA454424, AA180651,



AA842423, AA257507, AI063375, AA180692, AA051990, W15094, AA180620, W00308, AA406729, AA257712, AA257445, AA433148, C48534, AA495533, C08939, AA406845, AA933362, AA509242, AA842023, W29144, AI087490, AA842237, AA601823, W06538, AI087739, AI087524, AA842164, AA842642, AA471553, AA842419, AI113700, AA675813, AI083003, AI021727, AI083274, AI083309, W32823, AA253962, AA114520, AA430792, C47230, AA842674, AI053141, D75937, AA843034, W68979, AI105681, W91819, AA756971, AA509104, AA111828, AI109720, AI110161, AA933311, AA180602, AA842093, R86419, AI058057, AA186285, AA406891, AI105725, AA406888, AA649397, N94700, W51718, AA547916, AA840646, U74116, AA257756, AI082996, W91818, N74830, AA123585, AA118224, AA454446, AA841367, AA509109, AA842387, AI052833, AA751998, AA123634, AA257327, AA803962, AA842493, AA471448, AA406716, AA841361, AA840661, AA713447, AA406754, AA257695, AA841403, AA751834, AA406947, AI096182, AI105734, W15132, W59918, AA089352, AA180566, AA257427, AA257522, AA406980, AA471469, AA509264, AA509032, AA749469, AA752028, AA754167, AA754646, AA842335, AA842574, AI065957, AI065970, AI096289, AA109379, AA109472, AA114484, AA161689, AA430830, AA406948, AA433246, AA454371, AA471492, AA471703, AA752086, AA842624, AI096185, R47079, R86415, AA161711, AA161655, AA180549, AA257424, AA257437, AA257749, AA433393, AA471419, AA471602, AA430922, AA661401, AA752034, AA752066, AA842384, AI082936, AI083329, AI105677, Z33912, AA109362, AA109417, AA051807, AA430797, AA406799, AA430806, AA495548, AA661116, AI082951, AI105685, AI114069, R47062, W59884, AA109292, AA751829, AA754172, AA842257, AI082934, AI082958, AA180706, AA471516, AA508962, AA751563, AA751816, AA753137, AA753167, AA840972, AA842079, AA874756, AI105522, AI108824, AA180588, AA406676, AA471392, AA509142, AA840909, AA842175, AI066854, R47172, AA161565, AA257643, AA753161, AA471447, AA680450, AA752897, AA753237, AA842230, AA842544, AA161635, AA454428, AA508933, AA509309, C41215, AA842585, D73786, D75990, D75808, AA627049, N74809, AA257682, AA257517, AA257668, AA471470, AA509251, AA509051, AA990913, C39627, C41200, C42101, C46071, C48450, AA840970, AA842156, AA430902, AA842216, AA007706, AA406765, AA180574, AA842310, AI067585, AI077003, AA756933, AA114501, AA406690, AA454440, AA661371, AA056797, AA114372, AA180676, AI082955, AA406735, AA675746, AA752005, W29142, AA109261, AA842401, AA842602, AA246079, AI043420, AA626993, AA123655, AA257716, AA406931, AA791379, AI105668, AA842538, R47669, AA088150, AA109308, AA738555, AA123638, AA180582, AA123597, AA471535, AA180728, AA161741, AA186201, AA430817, AA433170, AA454407, AA471686, AA161596, AA509110, AA675874, AA471540, Z26577, AA753093, AI066829, D37716, AA627017, N99292, AA508936, AA257656, AA406768, AA471587, H39287, AA842431, AA161715, AA509204, AA979935

SEQ ID NO.: 2, AF086243, AE001154, X62889, , W67765, W67764, AA947751, AI141491, W76469, AA906091, AA872676, AA349825, H77545, H91001, W72232, AA948309, AA361403, R57582, AA337188, AA215714, AA481093, T75310, AA976452, , AI120744, AA462558, AI158491, AA014020, AA285990, AA051044, AA473453, AA896862, W56907, AA218305, AA020167, W85164, AA017810, W30604, AA541978, W20894, AA023164, AA760425, AA024084, AA840044, , AA231755, T20644, C45833, T24185, C68940, D37618, AA964657, AI011924

SEQ ID NO.: 3, L08501, Z97205, Z48950, AF010400, AC005162, L08502, D50608, , AA707653, AA861639, AA292496, AA702524, AI097367, AI138504, AI147933, AI141836, AI075247, N63868, W88668, AA703146, AA625621, AA292247, AI032848, AA005331, AA699781, AA427941, D31111, D31113, N92013, AA884207, AA044752, R70900, R16693, AA906542, R35112, AA481286, D80100, AA235512, AA960995, AA867982, D59403, W88874, AA558590, R10048, AA719917, H43573, R49500, AA744780, AA047136, AA789101, AA906332, AA747301, AA830606, AA434559, AA236698, AA328889, N98469, W17299, W46605, AA258082, AI078045, R11930, W74577, AA703312, AI082727, R45189, N76461, AA010500, W63646, W38891, AA490651, AA558805, W87891, AA160849, AA618177, AA776126, AA161281, , AA822308, AA499768, AA028780, AA183100, AA276783, AA120227, AA200285, AA212541, AA116265, AA821616, AA265629, AA108594, W61565, AA030311, W42275, AA419903, AA268027, AA028211, AA518504, AA125394, , C92235, C92002, C29917, AI145662, C65317, AA901847, AI137443, Z14719, D35515, R90625, H76970,

SEQ ID NO.: 4, AL031178, Y11905, AF031904, M34309, M29366, L33953, L33952,

L33956, L33954, L33957, U41289, X13369, AL022072, AJ223074, AF078695, AC004683, X17267, AF025526, AF071798, AC005274, AB004537, X68248, AF035537, AF056116, AB009052, AL023286, AF058701, L20725, AF001308, U22438, , AA211485, AA579574, AI078750, AA568661, AA604128, H49462, AA767424, H97012, AA565823, H49463, AA827171, T87152, F22114, AA748475, T31504, Z42997, T89930, R02581, H75949, T87057, AA322268, R02700, AA005034, AA576177, AI014302, AA348159, T83607, AA322497, AA211532, AA019517, R27657, AA026869, N67589, T83782, AA129383, Z39829, AA814308, AA425564, H67997, AA004420, AA644513, AA007691, T62593, AA333601, H75334, AI057250, T62521, N73865, AA004558, AA004484, AA007690, AA910241, AA456251, AA004568, AA341266, T28757, AA132360, AA781316, AA456942, AA782765, AA662593, AA461351, AA885220, N33840, AA843737, T62232, AA608559, AA643270, R44662, AI074863, AI051088, N63305, W24955, AA085886, W31918, AA047466, AA178965, AA369890, AA515015, N79466, AA768162, AA864694, AA811390, AA834531, AA293263, AA768335, AA749083, T57520, AA292001, AI085512, AA969032, AI027062, AA595663, AA827242, AA744475, , AA472485, AA097011, AA959170, AA647546, AA986669, AI006628, W12604, AI122356, AA571721, AA433697, AA880171, AA015463, AA116290, C87721, AA437983, AA990395, AA265925, W14785, W89262, AA059703, AA413613, AI020231, AA667024, AA407526, AA221491, AA222523, C76436, AA409700, W35766, , AA859485, H33704, T46512, C25899, T41882, C26586, C92302, AA660448, C94493, C25527, D35045, D33727, F14543, D37473, T45969, AA040979, D68274, T38420, AA228607, AA660541, T38283, AA842575, C71889, T38745, D68465, T46732, AA598393, C23775, C64798, AA042714, F19972, C65873, N97974, AA114425, AA900456

SEQ ID NO.: 5, U29344, S80437, M76767, X62889, M84761, J03514, X62888, X13415, X13135, U05714, AC004013, U58675, S47635, Z81533, AC005250, AC003661, AB008567, Y13444, X96401, AF026487, AF026488, , AA781445, AI037943, AI129371, C15883, AI073336, AA904077, AA569042, AI039428, AA568701, AA058907, AA911112, AA234022, N95359, AA565390, AA082427, AA588430, D60358, AA045488, AA084417, N72089, AI143390, R51974, AA081439, AA635907, T47621, AA579930, AA995057, AA827039, AA872490, AA588319, AA069032, AA062768, AA534011, W96404, H06082, AA938900, AA971262, AA836547, AA251544, AA250742, AA830405, AA906492, AA102653, R38286, AA258075, H83302, H38522, R61787, R33742, R24094, AI052406, AA931452, R23634, R39640, AA974568, AI028383, AA312451, AA148800, AI138982, R36172, AA926921, , AA472563, AA914598, AA423256, AA476186, AA041992, W84938, AA797706, AA709956, AA544361, AI006075, AA717202, AA008602, AA033399, AA472306, AI037430, AI152943, AA475582, AA471742, AA530292, AA277496, AA718588, AA823112, AA822771, AA710973, AA981780, AA718846, AA797557, AA510746, AI036049, AA543891, AA606513, AA981758, AA822268, AA413161, AA472201, AA646527, AA213036, AA168254, AA718549, AA458342, AA119418, AA879925, AA050917, AA718814, AA793638, AA177849, AA575627, AA537367, AI121788, AA203946, AA413157, AA517432, AA867736, AI037420, W30436, , Z71851, AA957415, AI112847, AA955881, AA963915, AI043663, C95061, AI009894, AA957229, AI044678

SEQ ID NO.: 6, AC002426, AC004674, AC004675, AC004006, L12157, U20839, U46028, U20835, AC004058, X67115, X76266, M25485, U59759, AC005172, U70848, AF042274, X14724, L31840, L41679, , W52480, AA863014, W56770, AA286755, AA164604, W52777, AA814246, AA765427, AA873647, AA360577, AI139274, AA770312, AA732557, AA568651, AA164603, W56724, AA194905, AA865009, R16884, R08110, F07672, AA577790, T11467, AA502489, AA366688, AA480628, AA516318, AA090005, AA081908, F01265, AA780686, , AA466811, AA465808, AA197646, AI157174, AA153086, AA197667, AA066612, AA058086, AA024186, AA110104, AA880419, AA870161, AA153880, AA072995, , AI031052, AI043934, AA849320, D15181, U95093, AI096188, T09655, C62926, AA996434, AA925202, AA817271, AA901051, R03660, AA948980, AA658709, AA899446, AA957150, AA800160, T76196, AA891787

SEQ ID NO.: 7, X63547, X63546, AF017306, U44839, U20657, AF017305, AJ001589, L21998, D21270, Z48245, U63834, Z68006, L04573, D63819, AC005266, AF025468, M94131, , AI056961, C15588, AA373847, AA887911, N87070, R39133, AA226825, AA353972, AA325352, R88378, AA490675, AA456219, AA701415, AA046611, AA456224, , AA065652, AA221447, AA710221, AA717401, AA254049,

AA041745, AA510261, AA170745, AA656404, AA572506, AA174539, AA217630, ,  
AI134284, C58457, C11202, AA413362, AA728532

SE ID NO.: 8, AB016886, Z92811, AL021481, U52078, L46702, U23515, AC005555,  
Z84814, Z93016, D21138, X57513, L07144, D83711, M28488, U15974, U52513,  
AF026939, L19120, J05258, Z13985, Z70691, U59227, , AA190526, AA764854,  
AI014655, H88220, AA622877, N32046, H89609, AA682362, AA464420, AA375477,  
AA284905, AA257109, N80276, H89373, N92393, AA884334, AA778708, AA766209,  
AA535677, W46414, AA770266, AA983635, AA721113, H88219, R87349, AA628091, ,  
AI116513, AA183589, AI035517, AA209952, AA184622, AI097904, AA016440,  
AA162370, AA939521, AA162376, AA718152, AA048154, AA253815, AA795350, ,  
AA996981, C44448, C41783, C44695, C44039, C62421, C47063, C45014, D36613,  
C60872, C50136, C69177, AI055708, D27648, C65648, AI114022, AI064663,  
AI107213, AA395461, AA800126, AA891487, T01611, C27943, AI113482, C83830

SEQ ID NO.: 9, L43510, U71249, L11275, X73541, Z28317, AB008270, X75652,  
U05987, U85262, U24189, Z46676, X70823, AC004981, D11079, X56121, M58053,  
U33636, Z92773, U23168, D31662, , R15557, F01629, N71722, AA252548,  
AA806751, T59557, AA612671, AA329585, AA295675, AA166990, AA128100, H46363,  
AA125810, , AA396888, AA408999, AA270873, AA144722, AA863954, W10303, ,  
C38016, N96746, Z37622, AA395862, C38756, F15295, N99294, W68877, C44749,  
C24349, F15569, AA113611, AA689147, N96676, C73598, F15564, L46559, W06235,  
T09718, H77154, T14151, AI082914, AA650815, AA728021, AA072559, C92112,  
T46743, C61743, AA550223, C12798, AA275531, AA681003, T75878, AI100047,  
T75882, AA848187, T44109, AA542686, H21339

SEQ ID NO.: 10, Z81595, U41372, U01317, AE000696, AC002057, X66250, L11665,  
D13438, U29377, Z50028, U40953, AA114228, AI025080, R80188, W28745,  
AA381819, AA381991, Z44165, H75915, AI124743, H08139, AA375957, AA381750,  
AA166751, R42511, T26985, , AA471592, AA661387, AA606231, D35147

SEQ ID NO.: 11, X76498, Z97632, AC004232, AC005184, AC002545, AC003991.

### EQUIVALENTS

Those skilled in the art will recognize, or be able to ascertain using no more than routine experimentation, many equivalents to the specific embodiments of the invention described herein. Such equivalents are intended to be encompassed by the following claims.

All references disclosed herein are incorporated by reference in their entirety.

We claim:

Claims

1. A method of diagnosing a disorder characterized by expression of a human cancer associated antigen precursor coded for by a nucleic acid molecule, comprising:  
contacting a biological sample isolated from a subject with an agent that specifically binds to the nucleic acid molecule, an expression product thereof, or a fragment of an expression product thereof complexed with an HLA molecule, wherein the nucleic acid molecule is a NA Group 1 nucleic acid molecule, and  
determining the interaction between the agent and the nucleic acid molecule or the expression product as a determination of the disorder.
2. The method of claim 1, wherein the agent is selected from the group consisting of
  - (a) a nucleic acid molecule comprising NA group 1 nucleic acid molecules or a fragment thereof,
  - (b) a nucleic acid molecule comprising NA group 3 nucleic acid molecules or a fragment thereof,
  - (c) a nucleic acid molecule comprising NA group 5 nucleic acid molecules or a fragment thereof,
  - (d) an antibody that binds to an expression product of NA group 1 nucleic acids,
  - (e) an antibody that binds to an expression product of NA group 3 nucleic acids,
  - (f) an antibody that binds to an expression product of NA group 5 nucleic acids,
  - (g) an agent that binds to a complex of an HLA molecule and a fragment of an expression product of a NA group 1 nucleic acid,
  - (h) an agent that binds to a complex of an HLA molecule and a fragment of an expression product of a NA group 3 nucleic acid, and
  - (i) an agent that binds to a complex of an HLA molecule and a fragment of an expression product of a NA group 5 nucleic acid.
3. The method of claim 1, wherein the disorder is characterized by expression of a plurality of human cancer associated antigen precursors and wherein the agent is a plurality of agents, each of which is specific for a different human cancer associated antigen precursor, and wherein said plurality of agents is at least 2, at least 3, at least 4, at least 4, at least 6, at least 7, or at least 8, at least 9 or at least 10 such agents.

4. The method of claims 1-3, wherein the agent is specific for a human cancer associated antigen precursor that is selected from the group consisting of breast, gastric, lung, prostate, renal, colon, thyroid, Hodgkin's disease, and hepatocarcinoma cancer associated antigen precursors.
5. A method for determining regression, progression or onset of a condition characterized by expression of abnormal levels of a protein encoded by a nucleic acid molecule that is a NA Group 1 molecule, comprising
  - monitoring a sample, from a patient who has or is suspected of having the condition, for a parameter selected from the group consisting of
    - (i) the protein,
    - (ii) a peptide derived from the protein,
    - (iii) an antibody which selectively binds the protein or peptide, and
    - (iv) cytolytic T cells specific for a complex of the peptide derived from the protein and an MHC molecule,as a determination of regression, progression or onset of said condition.
6. The method of claim 5, wherein the sample is a body fluid, a body effusion or a tissue.
7. The method of claim 5, wherein the step of monitoring comprises contacting the sample with a detectable agent selected from the group consisting of
  - (a) an antibody which selectively binds the protein of (i), or the peptide of (ii),
  - (b) a protein or peptide which binds the antibody of (iii), and
  - (c) a cell which presents the complex of the peptide and MHC molecule of (iv).
8. The method of claim 7, wherein the antibody, the protein, the peptide or the cell is labeled with a radioactive label or an enzyme.
9. The method of claim 5, comprising assaying the sample for the peptide.
10. The method of claim 5, wherein the nucleic acid molecule is a NA Group 3 molecule.

11. The method of claim 5, wherein the nucleic acid molecule is a NA Group 5 molecule.
12. The method of claim 5, wherein the protein is a plurality of proteins, the parameter is a plurality of parameters, each of the plurality of parameters being specific for a different of the plurality of proteins, at least one of which is a cancer associated protein encoded by a NA Group 1 molecule.
13. The method of claim 5, wherein the protein is a plurality of proteins, at least one of which is kinectin, the remainder of which are non-kinectin cancer associated proteins, and wherein the parameter is a plurality of parameters, each of the plurality of parameters being specific for a different of the plurality of proteins.
14. A pharmaceutical preparation for a human subject comprising  
an agent which when administered to the subject enriches selectively the presence of complexes of an HLA molecule and a human cancer associated antigen, and  
a pharmaceutically acceptable carrier, wherein the human cancer associated antigen is a fragment of a human cancer associated antigen precursor encoded by a nucleic acid molecule comprises a NA Group 1 molecule.
15. The pharmaceutical preparation of claim 14, wherein the agent comprises a plurality of agents, each of which enriches selectively in the subject complexes of an HLA molecule and a different human cancer associated antigen, wherein at least one of the human cancer associated antigens is encoded by a NA Group 1 molecule.
16. The pharmaceutical preparation of claim 15, wherein the plurality is at least two, at least three, at least four or at least 5 different such agents.
17. The pharmaceutical preparation of claim 14, wherein the nucleic acid molecule is a NA Group 3 nucleic acid molecule.
18. The pharmaceutical preparation of claim 14, wherein the agent comprises a plurality of

agents, at least one of which is kinectin, the remainder of which are non-kinectin cancer associated proteins, and each of which enriches selectively in the subject complexes of an HLA molecule and a different human cancer associated antigen.

19. The pharmaceutical preparation of claim 14, wherein the agent is selected from the group consisting of

- (1) an isolated polypeptide comprising the human cancer associated antigen, or a functional variant thereof,
- (2) an isolated nucleic acid operably linked to a promoter for expressing the isolated polypeptide, or functional variant thereof,
- (3) a host cell expressing the isolated polypeptide, or functional variant thereof, and
- (4) isolated complexes of the polypeptide, or functional variant thereof, and an HLA molecule.

20. The pharmaceutical preparation of claims 14-19, further comprising an adjuvant.

21. The pharmaceutical preparation of claim 14, wherein the agent is a cell expressing an isolated polypeptide comprising the human cancer associated antigen or a functional variant thereof, and wherein the cell is nonproliferative.

22. The pharmaceutical preparation of claim 14, wherein the agent is a cell expressing an isolated polypeptide comprising the human cancer associated antigen or a functional variant thereof, and wherein the cell expresses an HLA molecule that binds the polypeptide.

23. The pharmaceutical preparation of claim 21 or 22, wherein the isolated polypeptide comprises a kinectin polypeptide.

24. The pharmaceutical preparation of claim 14, wherein the agent is at least two, at least three, at least four or at least five different polypeptides, each coding for a different human cancer associated antigen or functional variant thereof, wherein at least one of the human cancer associated antigens is encoded by a NA Group I molecule.

25. The pharmaceutical preparation of claim 24, wherein the at least one of the human cancer associated antigens is kinectin or a fragment thereof.
26. The pharmaceutical preparation of claim 14, wherein the agent is a PP Group 2 polypeptide.
27. The pharmaceutical preparation of claim 14, wherein the agent is a PP Group 3 polypeptide or a PP Group 4 polypeptide.
28. The pharmaceutical preparation of claim 22, wherein the cell expresses one or both of the polypeptide and HLA molecule recombinantly.
29. The pharmaceutical preparation of claim 22, wherein the cell is nonproliferative.
30. A composition comprising  
an isolated agent that binds selectively a PP Group 1 polypeptide.
31. The composition of matter of claim 30, wherein the agent binds selectively a PP Group 2 polypeptide.
32. The composition of matter of claim 30, wherein the agent binds selectively a PP Group 3 polypeptide.
33. The composition of matter of claim 30, wherein the agent binds selectively a PP Group 4 polypeptide.
34. The composition of matter of claim 30, wherein the agent binds selectively a PP Group 5 polypeptide.
35. The composition of claims 30-34, wherein the agent is a plurality of different agents that bind selectively at least two, at least three, at least four, or at least five different such polypeptides.



36. The composition of claim 35, wherein the at least one of polypeptides is kinectin or a fragment thereof.
37. The composition of claims 30-34, wherein the agent is an antibody.
38. The composition of claim 315 wherein the agent is an antibody.
39. A composition of matter comprising  
a conjugate of the agent of claims 30-34 and a therapeutic or diagnostic agent.
40. A composition of matter comprising  
a conjugate of the agent of claim 35 and a therapeutic or diagnostic agent.
41. The composition of matter of claim 39, wherein the conjugate is of the agent and a therapeutic or diagnostic that is a toxin.
42. A pharmaceutical composition comprising an isolated nucleic acid molecule selected from the group consisting of NA Group 1 molecules and NA Group 2 molecules, and a pharmaceutically acceptable carrier.
43. The pharmaceutical composition of claim 42, wherein the isolated nucleic acid molecule comprises a NA Group 3 or NA Group 4 molecule.
44. The pharmaceutical composition of claim 42, wherein the isolated nucleic acid molecule comprises at least two isolated nucleic acid molecules coding for two different polypeptides, each polypeptide comprising a different human cancer associated antigen.
45. The pharmaceutical composition of claim 44, wherein at least one of the polypeptides is a kinectin polypeptide.
46. The pharmaceutical composition of claims 42-45 further comprising an expression

vector with a promoter operably linked to the isolated nucleic acid molecule.

47. The pharmaceutical composition of claims 42-45 further comprising a host cell recombinantly expressing the isolated nucleic acid molecule.
48. A pharmaceutical composition comprising  
an isolated polypeptide comprising a PP Group 1 or a PP Group 2 polypeptide, and  
a pharmaceutically acceptable carrier.
49. The pharmaceutical composition of claim 48, wherein the isolated polypeptide comprises a PP Group 3 or a PP Group 4 polypeptide.
50. The pharmaceutical composition of claim 48, wherein the isolated polypeptide comprises at least two different polypeptides, each comprising a different human cancer associated antigen.
51. The pharmaceutical composition of claim 50, wherein at least one human cancer associated antigen is kinectin.
52. The pharmaceutical composition of claim 48, wherein the isolated polypeptides are PP Group 11 polypeptides or HLA binding fragments thereof.
53. The pharmaceutical composition of claim 48, wherein the isolated polypeptides are PP Group 12 polypeptides or HLA binding fragments thereof.
54. The pharmaceutical composition of claims 48-53, further comprising an adjuvant.
55. An isolated nucleic acid molecule comprising a NA Group 3 molecule.
56. An isolated nucleic acid molecule comprising a NA Group 4 molecule.
57. An isolated nucleic acid molecule selected from the group consisting of

(a) a fragment of a nucleic acid molecule having a nucleotide sequence selected from the group consisting of nucleotide sequences set forth as SEQ ID NOs. 1-11 and 22-46, of sufficient length to represent a sequence unique within the human genome, and identifying a nucleic acid encoding a human cancer associated antigen precursor,

(b) complements of (a),

provided that the fragment includes a sequence of contiguous nucleotides which is not identical to any sequence selected from the sequence group consisting of

(1) sequences having the GenBank accession numbers of Table 1,

(2) complements of (1), and

(3) fragments of (1) and (2).

58. The isolated nucleic acid molecule of claim 50, wherein the sequence of contiguous nucleotides is selected from the group consisting of:

- (1) at least two contiguous nucleotides nonidentical to the sequence group,
- (2) at least three contiguous nucleotides nonidentical to the sequence group,
- (3) at least four contiguous nucleotides nonidentical to the sequence group,
- (4) at least five contiguous nucleotides nonidentical to the sequence group,
- (5) at least six contiguous nucleotides nonidentical to the sequence group,
- (6) at least seven contiguous nucleotides nonidentical to the sequence group.

59. The isolated nucleic acid molecule of claim 57, wherein the fragment has a size selected from the group consisting of at least: 8 nucleotides, 10 nucleotides, 12 nucleotides, 14 nucleotides, 16 nucleotides, 18 nucleotides, 20, nucleotides, 22 nucleotides, 24 nucleotides, 26 nucleotides, 28 nucleotides, 30 nucleotides, 50 nucleotides, 75 nucleotides, 100 nucleotides, and 200 nucleotides.

60. The isolated nucleic acid molecule of claim 57, wherein the molecule encodes a polypeptide which, or a fragment of which, binds a human HLA receptor or a human antibody.

61. An expression vector comprising an isolated nucleic acid molecule of any of claims 55-60 operably linked to a promoter.

62. An expression vector comprising a nucleic acid operably linked to a promoter, wherein the nucleic acid is a NA Group 2 molecule.
63. An expression vector comprising a NA Group 1 or Group 2 molecule and a nucleic acid encoding an HLA molecule.
64. A host cell transformed or transfected with an expression vector of claim 61.
65. A host cell transformed or transfected with an expression vector of claims 62 or 63.
66. A host cell transformed or transfected with an expression vector of claim 61 and further comprising a nucleic acid encoding HLA.
67. A host cell transformed or transfected with an expression vector of claim 62 and further comprising a nucleic acid encoding HLA.
68. An isolated polypeptide encoded by the isolated nucleic acid molecule of claim 55 or claim 56.
69. A fragment of the polypeptide of claim 68 which is immunogenic.
70. An isolated polypeptide comprising a fragment of a kinectin polypeptide which is immunogenic.
71. The fragment of claims 69 or 70, wherein the fragment, or a portion of the fragment, binds HLA or a human antibody.
72. An isolated fragment of a human cancer associated antigen precursor which, or portion of which, binds HLA or a human antibody, wherein the precursor is encoded by a nucleic acid molecule that is a NA Group 1 molecule.

73. The fragment of claim 72, wherein the fragment is part of a complex with HLA.
74. The fragment of claim 73, wherein the fragment is between 8 and 12 amino acids in length.
75. An isolated polypeptide comprising a fragment of the polypeptide of claim 68 of sufficient length to represent a sequence unique within the human genome and identifying a polypeptide that is a human cancer associated antigen precursor.
76. A kit for detecting the presence of the expression of a human cancer associated antigen precursor comprising  
a pair of isolated nucleic acid molecules each of which consists essentially of a molecule selected from the group consisting of (a) a 12-32 nucleotide contiguous segment of the nucleotide sequence of any of the NA Group 1 molecules and (b) complements of ("a"), wherein the contiguous segments are nonoverlapping.
77. The kit of claim 76, wherein the pair of isolated nucleic acid molecules is constructed and arranged to selectively amplify an isolated nucleic acid molecule that is a NA Group 3 molecule.
78. A method for treating a subject with a disorder characterized by expression of a human cancer associated antigen precursor, comprising  
administering to the subject an amount of an agent, which enriches selectively in the subject the presence of complexes of an HLA molecule and a human cancer associated antigen, effective to ameliorate the disorder, wherein the human cancer associated antigen is a fragment of a human cancer associated antigen precursor encoded by a nucleic acid molecule selected from the group consisting of  
(a) a nucleic acid molecule comprising NA group 1 nucleic acid molecules,  
(b) a nucleic acid molecule comprising NA group 3 nucleic acid molecules,  
(c) a nucleic acid molecule comprising NA group 5 nucleic acid molecules.
79. The method of claim 78, wherein the disorder is characterized by expression of a

plurality of human cancer associated antigen precursors and wherein the agent is a plurality of agents, each of which enriches selectively in the subject the presence of complexes of an HLA molecule and a different human cancer associated antigen, wherein at least one of the human cancer associated antigens is encoded by a NA Group 1 molecule.

80. The method of claim 79, wherein at least one of the human cancer associated antigens is kinectin or a fragment thereof.

81. The method of claim 79, wherein the plurality is at least 2, at least 3, at least 4, or at least 5 such agents.

82. The method of claims 78-81, wherein the agent is an isolated polypeptide selected from the group consisting of PP Group 1, PP Group 2, PP Group 3, PP Group 4, and PP Group 5.

83. The method of claims 78-81, wherein the disorder is cancer.

84. The method of claims 82, wherein the disorder is cancer.

85. A method for treating a subject having a condition characterized by expression of a human cancer associated antigen precursor in cells of the subject, comprising:

- (i) removing an immunoreactive cell containing sample from the subject,
- (ii) contacting the immunoreactive cell containing sample to the host cell under conditions favoring production of cytolytic T cells against a human cancer associated antigen which is a fragment of the precursor,
- (iii) introducing the cytolytic T cells to the subject in an amount effective to lyse cells which express the human cancer associated antigen, wherein the host cell is transformed or transfected with an expression vector comprising an isolated nucleic acid molecule operably linked to a promoter, the isolated nucleic acid molecule being selected from the group of nucleic acid molecules consisting of NA Group 1, NA Group 2, NA Group 3, NA Group 4, and NA Group 5.

86. The method of claim 85, wherein the host cell recombinantly expresses an HLA molecule which binds the human cancer associated antigen.
87. The method of claim 85, wherein the host cell endogenously expresses an HLA molecule which binds the human cancer associated antigen.
88. A method for treating a subject having a condition characterized by expression of a human cancer associated antigen precursor in cells of the subject, comprising:
- (i) identifying a nucleic acid molecule expressed by the cells associated with said condition, wherein said nucleic acid molecule is a NA Group 1 molecule;
  - (ii) transfecting a host cell with a nucleic acid selected from the group consisting of (a) the nucleic acid molecule identified, (b) a fragment of the nucleic acid identified which includes a segment coding for a human cancer associated antigen, (c) deletions, substitutions or additions to (a) or (b), and (d) degenerates of (a), (b), or (c);
  - (iii) culturing said transfected host cells to express the transfected nucleic acid molecule, and;
  - (iv) introducing an amount of said host cells or an extract thereof to the subject effective to increase an immune response against the cells of the subject associated with the condition.
89. The method of claim 88, wherein the nucleic acid molecule is a kinectin nucleic acid molecule.
90. The method of claim 88, further comprising identifying an MHC molecule which presents a portion of an expression product of the nucleic acid molecule, wherein the host cell expresses the same MHC molecule as identified and wherein the host cell presents an MHC binding portion of the expression product of the nucleic acid molecule.
91. The method of claim 88, wherein the immune response comprises a B-cell response or a T cell response.
92. The method of claim 91, wherein the response is a T-cell response which comprises

generation of cytolytic T-cells specific for the host cells presenting the portion of the expression product of the nucleic acid molecule or cells of the subject expressing the human cancer associated antigen.

93. The method of claim 88, wherein the nucleic acid molecule is a NA Group 3 molecule.
94. The method of claims 88 or 90, further comprising treating the host cells to render them non-proliferative.
95. A method for treating or diagnosing or monitoring a subject having a condition characterized by expression of an abnormal amount of a protein encoded by a nucleic acid molecule that is a NA Group 1 molecule, comprising  
administering to the subject an antibody which specifically binds to the protein or a peptide derived therefrom, the antibody being coupled to a therapeutically useful agent, in an amount effective to treat the condition.
96. The method of claim 95, wherein the antibody is a monoclonal antibody.
97. The method of claim 96, wherein the monoclonal antibody is a chimeric antibody or a humanized antibody.
98. A method for treating a condition characterized by expression in a subject of abnormal amounts of a protein encoded by a nucleic acid molecule that is a NA Group 1 nucleic acid molecule, comprising  
administering to a subject a pharmaceutical composition of any one of claims 14-29 and 42-547 in an amount effective to prevent, delay the onset of, or inhibit the condition in the subject.
99. The method of claim 98, wherein the condition is cancer.
100. The method of claim 98, further comprising first identifying that the subject expresses



in a tissue abnormal amounts of the protein.

101. The method of claim 99, further comprising first identifying that the subject expresses in a tissue abnormal amounts of the protein.

102. A method for treating a subject having a condition characterized by expression of abnormal amounts of a protein encoded by a nucleic acid molecule that is a NA Group 1 nucleic acid molecule, comprising

- (i) identifying cells from the subject which express abnormal amounts of the protein;
- (ii) isolating a sample of the cells;
- (iii) cultivating the cells, and
- (iv) introducing the cells to the subject in an amount effective to provoke an immune response against the cells.

103. The method of claim 102, further comprising rendering the cells non-proliferative, prior to introducing them to the subject.

104. A method for treating a pathological cell condition characterized by aberrant expression of a protein encoded by a nucleic acid molecule that is a NA Group 1 nucleic acid molecule, comprising

administering to a subject in need thereof an effective amount of an agent which inhibits the expression or activity of the protein.

105. The method of claim 104, wherein the agent is an inhibiting antibody which selectively binds to the protein and wherein the antibody is a monoclonal antibody, a chimeric antibody, a humanized antibody or an antibody fragment.

106. The method of claim 104, wherein the agent is an antisense nucleic acid molecule which selectively binds to the nucleic acid molecule which encodes the protein.

107. The method of claim 104, wherein the nucleic acid molecule is a NA Group 3 nucleic acid molecule.

108. The method of claim 104, wherein the nucleic acid molecule is a kinectin nucleic acid molecule.
109. A composition of matter useful in stimulating an immune response to a plurality of a proteins encoded by nucleic acid molecules that are NA Group 1 molecules, comprising  
a plurality of peptides derived from the amino acid sequences of the proteins, wherein the peptides bind to one or more MHC molecules presented on the surface of the cells which express an abnormal amount of the protein.
110. The composition of matter of claim 109, wherein at least a portion of the plurality of peptides bind to MHC molecules and elicit a cytolytic response thereto.
111. The composition of matter of claim 109, wherein at least one of the proteins is kinectin.
112. The composition of matter of claim 110, further comprising an adjuvant.
113. The composition of matter of claim 112, wherein said adjuvant is a saponin, GM-CSF, or an interleukin.
114. The composition of matter of claim 109, further comprising at least one peptide useful in stimulating an immune response to at least one protein which is not encoded by nucleic acid molecules that are NA Group 1 molecules, wherein the at least one peptide binds to one or more MHC molecules.
115. An isolated antibody which selectively binds to a complex of:
- (i) a peptide derived from a protein encoded by a nucleic acid molecule that is a NA Group 1 molecule and
  - (ii) and an MHC molecule to which binds the peptide to form the complex,
- wherein the isolated antibody does not bind to (i) or (ii) alone.

116. The antibody of claim 115, wherein the antibody is a monoclonal antibody, a chimeric antibody, a humanized antibody, or a fragment thereof.

## SEQUENCE LISTING

<110> Ludwig Institute for Cancer Research

5 <120> Cancer Associated Antigens  
and Uses Therefor

<130> L0461/7050WO

10 <140> US 09/166,300  
<141> 1997-10-05

<140> US 09/166,350  
15 <141> 1997-10-05

<160> 45

<170> FastSEQ for Windows Version 3.0

20 <210> 1  
<211> 4422  
<212> DNA  
<213> Homo sapiens

25 <400> 1

gattattctg	gaggaagatg	acottcattt	tatggttatg	tgtgtgtgtc	tgtgtgtctt	60
tggaataata	tatataattt	tttcaaatag	gaagccaaca	tatcaagtga	tgaattaaag	120
tatgctgcag	aatatatatt	ctaaaactac	aaaaaaagta	ctgaatatca	aaatgataca	180
gcttatacat	atagttactg	tgacaagtga	cagactgcta	gttcagaatt	caaaaatcct	240
ttcctagttt	gtgagataat	gggctaattt	ccttctgcct	gccactgggg	caaagcaaat	300
tgctttagtt	ttttagtgaga	gttcttagaa	gtttgttggt	attccttcat	ccacagcatc	360
cattgttgaa	ataaccattt	tcagttgtga	tgccctaact	aagaagccaa	ttgttagcct	420
gaaatgcaat	cttggttagcc	agtttcaatg	aagctagaga	ttagtcagaa	aaagttagct	480
gttgggcttt	agaaagggat	tttgagtcct	gncatttcta	cttgggagca	ttttggagca	540
gattaaagctt	tcagtataaaa	aacaagtggc	tacctgatgg	aaacttttct	tacctttata	600
gggaaactga	gcacaagctg	aatgatattg	tctgtctgcaa	aaaaaaacaa	acaaaaaaaa	660
aaacaaaacaa	aaaacaaaaa	aaaaaaataa	aaaacctcgt	gccgaattcg	gcacgagggg	720
aagccccctg	cacccccctg	cctccggcgg	ccgcccggcc	gctggccctg	cagccgtcgc	780
cgctgcctcg	ggctacagcc	ccgggctcgg	cggtcccggc	tggggaagga	gggcccgcag	840
cgcgctccgga	gcccgcggag	atggcgggag	ggcactgcgg	cagcttcccc	gcggcggcgg	900
ccggcagcgg	cgagatcgtc	caactgaacg	taggggggac	cagatttagt	acctcaagac	960
aaactcttat	gtggattcca	gattcttttt	tttccagttt	gctgagtggg	agaatttcaa	1020
cacttcgaga	tgaaacttgt	gctatattta	ttgatagaga	tccagcagca	tttgcaccca	1080
ttttaaattt	tcttcggaca	aaagaactag	acttaagggg	agtgagtatt	aatgtttcta	1140
ggcatgaagc	agaaattttac	gggatcactc	cattagtaag	aaggcttctc	ttatgtgaag	1200
aattggagcg	ttcctcttgt	ggcagtgctc	tttttcatgg	ttacttgccc	ccaccaggta	1260
ttcctagtcg	taaaataaac	aacacagtca	gatctgctga	ttctaggaat	ggtctaaatt	1320
ctacagaagg	tgaagcccgg	ggaaatggta	cacagcctgt	tctctctgga	acgggagaag	1380
aaactgttag	gctaggattt	cctgtggatc	cacgaaaggt	gctaatagta	gctggccatc	1440
acaactggat	tgtagctgca	tatgccattt	ttgctgngtg	gtacagaatc	aaagaatctt	1500
caggatggca	gcaagtgttt	acgagcccat	atttggattg	gactatcgaa	cgagtagctt	1560
taaatgcaaa	ggtggttgga	gggccacatg	gagacaaaga	caaaatggtt	gctgttgccct	1620
cagagagtag	catcatcttg	tggagtgttc	aggatggggg	aagtgggaag	gaaattggag	1680
tgttcagcct	gggtgttcct	gtagatgctc	tcttctttat	tggtaaccag	ttggtggcca	1740
cgagtcatac	agggaaagtg	ggagtgtgga	atgctgtcac	tcagcactgg	caggttcaag	1800
atgttgttcc	tataactagt	tatgacactg	ctggatcatt	ccttctgctt	ggatgtaaca	1860
atggatcaat	atattacata	gatatgcaga	agttcccctt	gcgaatgaaa	gataatgatc	1920
ttcttgtaac	tgaactgtat	catgatcctt	caaagtgatc	tattactgct	ctgagtgttt	1980
acctcacacc	caaaacaagt	gtcagtggtt	catagatcga	gatcgccat	ggtacgagct	2040
ctggagcagt	acgagtgttt	gtacaacacc	cagagacagt	tgggtcaggt	cctcagcttt	2100
ttcagacttt	cacagttcac	cgaagtcccg	taacaaaaat	catgctatca	gagaagcatc	2160
ttgtatcagt	ctgtgcagat	aataatcatg	tccggacgtg	gacagtaaca	cgattcngag	2220
gaatgatctc	tactcagcca	ggttctactc	cttttagcgtc	attcaagata	ctatccctgg	2280
aggagacaga	aagtcattgt	agctattcct	ctggaaatga	cataggacct	tttggagagc	2340
gagacgatca	acaggtgttt	atccagaaag	ttgttcccat	caccaacaaa	ctatttgtaa	2400
gactctcatc	gactggaaaa	agaatatgtg	agatccaggc	tgttgactgt	actacaatat	2460
cctcattttac	agtgagggaa	tgtgagggat	ccagtaggat	gggctcaaga	ccaaggcgcct	2520

	acttggtcac	aggccatata	aatggcagta	ttcaaatgtg	ggatctgacc	actgctatgg	2580
	atatgggttaa	caaaagtga	gataaggatg	taggtggtcc	aaccgaagaa	gagctactca	2640
	aattactcga	tcaatgtgat	ttgagcacat	ctcgctgtgc	tactcctaac	atcagtcacg	2700
	caacttccgt	agtttcagcat	agccacttac	gagaatcaaa	ttctagcctt	cagcttcagc	2760
5	accatgatac	cacccatgaa	gcagctactt	acggttccat	gaggccttac	agagaaagtc	2820
	ctttattagc	aagggcaaga	aggactgaga	gctttcacag	ttatagggac	ttccagacta	2880
	ttaatttgaa	cagaaatgta	gaaagagctg	tccctgaaaa	tggttaacttg	ggtccaatac	2940
	aagctgaagt	gaaaggggca	acaggggaat	gtaatatatc	tgagagaaag	tctcctggag	3000
	tagaaaataaa	aagtttgaga	gaattggata	gtggattgga	agtgcataaa	atagctgaag	3060
10	gtttttcaga	atccaagaaa	aggctcatcag	aagatgaaaa	tgaaaaataaa	atagagttaa	3120
	ggaagaaaag	aggatttgaa	gggggaggat	tccttggaag	aaagaaagtt	ccctatctgg	3180
	catcatcacc	aagtacttcc	gatggaggaa	ctgactcacc	tggtactgcg	tccccatctc	3240
	ctacaaagac	tactccatct	cctcggcata	aaaaaagtga	ttcttcaggt	caggagtaca	3300
	gcttgtagaa	actcaccaaa	atgaatagtt	gtttcgggta	catttagatg	aaagttaaac	3360
15	tttactgaat	ttcagtagat	tagtttttac	actaaaactt	tacaagataa	aattggactt	3420
	catttagtat	ctttttaaca	gaattacttg	gaataatgag	atacaataat	catatctctt	3480
	ttgacatttt	ggaaattttt	ttaattttac	aagtacattt	aacagatcat	ttataaagca	3540
	ggagtcattt	tttaacactta	ccgacttttt	ttggtttgga	aacatattac	cacgtcttaa	3600
	taggatgggtg	cccatatagg	tgagcatccc	tttagatcat	gggaaccagc	agactgcatt	3660
20	cctaattcttc	attatgcctg	agacttgtct	tacaatgtta	cctttaagtg	aatcacataa	3720
	ttgtcttttg	aacttggctt	cccaacactt	attgtgattg	caaagtgttt	accagatatt	3780
	tgatgaggtg	ctatgtttgt	gaaaaacata	tcattgtaatt	caaaaacact	attgatattg	3840
	aataccagat	accactatgt	agtaagtctt	ttaggatgat	tttaatttag	tcgtgcgtca	3900
	ttttctgatt	ctcatcattg	ggagatctta	aatcttagca	agcattagca	atattaaatg	3960
25	ccaaaattcc	attgaaactt	tcaagttgga	gcaattgtct	gtgtttgaaa	agatgaaata	4020
	aaaataataa	tcaagggcaa	agctttgagt	gccacgaagg	gaaagctgta	ccagttgcta	4080
	acctgtcttg	tttcaggagc	caccatgttt	ttttttcagt	gttancaaca	atcatgataa	4140
	ttaaattaaa	acnctagttt	gttcacttgt	aggactgcag	ttctgaattt	tgggttaaag	4200
	gttttggtctg	ctgtaagaat	gtgaatttga	atgtattttg	aattgtaaga	gcaaaagaac	4260
30	gtttttgtac	aatttttttt	cattttaattg	gaatgatctt	cagggtttcta	caaatagggt	4320
	aattgtaaat	ttaaagcatt	agcattttatt	ggtgaataat	gtatatatcc	ccattccaag	4380
	aaatataagt	gagtgaagtt	gaaataaaat	cttttaaaat	ta		4422
	<210> 2						
35	<211> 502						
	<212> DNA						
	<213> Homo sapiens						
	<400> 2						
40	gctagcttctg	gcgcggatcc	ctgggcgtcc	gtacgtcgga	gtccttcgtc	ctccagggtc	60
	cctgtttcttt	gcgccagcgg	gaaccactat	ctctgcactc	ctgggggtttt	gttacatggc	120
	tgcttttcctc	aaaatgagtg	ttagtgtcaa	tttcttcaga	cctttcacca	ggtttttggt	180
	gccattttacc	cttcatagga	agagaaataa	cttaacaatt	ttgcagagat	acatgtcttc	240
	caaaataacca	gctgttactt	atcctaataa	tgagagtaca	cgcccttctg	aagagctaga	300
45	gttggtgataa	tggaataacta	ccatgaaatc	tagtgtgcaa	gaagaatgtg	tttcaacaat	360
	ctcaagcagt	aaggatgaag	atcctctagc	tgccaccaga	gagttcattg	agatgtggag	420
	attgcttggtc	agagaagtac	cagaacacat	cactgaagaa	gagctcaaaa	cccttatgga	480
	atgtgtttct	aacacagcaa	aa				502
	<210> 3						
50	<211> 1948						
	<212> DNA						
	<213> Homo sapiens						
	<400> 3						
55	ccatgtgagg	gaggggcccg	ctcctgcttg	gtgacagagt	cagcacgcgg	tggcctgcag	60
	ttcctccagc	agtgtgaccg	ggaggtctctg	gtggaatttg	ctctgcctca	gctggctcag	120
	gttgtgaccg	tgtatgagtt	tcttctgatg	aagggtgaaa	aagatcatct	agcaaaagcct	180
	tttttcccag	ctatatataa	ggaatttgaa	gagttgcata	aaatgggtta	gaaaatgtgc	240
60	caagattacc	tcagtatgtc	tggtctgtgt	tcccaggaga	ccctggaat	aaacaatgat	300
	aaggttgctg	agtcattagg	aatcacagaa	ttcctacgga	agaaagaaat	acaccagac	360
	aaccttgtag	ccaagcacct	cagccgagac	atggatggg	agcagctaga	gggagctagc	420
	agcgagaaga	gggaacgtga	ggctgcggag	gagggactgg	cctcagtga	aaggcccaga	480
	agagaagccc	tgtccaacga	taccactgaa	tctcttgctg	ccaacagcag	aggccgggag	540
65	aagcccaggc	ccttgcatgc	tttgcctgct	ggtttttccc	ctccagtaaa	tgtgactgtc	600
	tctcccgtt	ctgaagaag	ccatacaacg	acggtttctg	gtggcaatgg	gagcgtgttc	660
	caggcgggccc	cgcagcttca	ggcactggct	aacttagaag	ccaggagggg	gtctataggt	720

	gctgctctct	catcccggga	tgtcagtg	ctgctgttt	atgctcagtc	aggagagcct	780
	aggaggctga	cccaggcaca	ggtggcagcg	tttctggag	agaatgcttt	ggaacactct	840
	tcagaccagg	acacctggga	cagcctgagg	agcccgggtt	tctgcagccc	tttgtcatct	900
5	ggtgggtggag	cagagtcctt	gccgcctggg	gggcctggac	atgcagaggc	aggacacctc	960
	ggcaagggtt	gtgacttcca	cctgaaccac	cagcagccca	gccccaccag	cgctctgcct	1020
	acagagggtg	cagcccctcc	gcttgagaaa	atcttgtctg	tggatagcgt	ggcagtgga	1080
	tgtgcctaca	ggactgtgcc	caagccaggg	cctcagcctg	gcccacatgg	atcactattg	1140
	actgaagggt	gtctcagaag	cctttcgggg	gacttgaacc	ggttcccctg	tgggatggag	1200
	gtgcactctg	gccagagaga	actggagagc	gtggttgctg	tcggcgaagc	catggccttg	1260
10	aaatttccaa	tgaggagccat	gagttactgt	ctcagggaca	gaagcagatt	tttattcaga	1320
	cttccgatgg	gcttatcttg	tcccctccag	gtacaatagt	gtctcaggag	gaggacattg	1380
	tcacagtac	tgatgcagag	gggcgtgcct	gcgatgggc	ccgctagaag	gagttcctct	1440
	agaagctgtg	gagtcggtcg	tcaccgcgag	agccctcaca	gtgaagtggg	gtcagatcct	1500
	agattcgtct	gattttatcc	agagaagggtc	tctgggaagc	aatgtatatt	tttctaattg	1560
15	gaatattgca	cagatgaacc	ttttatttat	aaagaataat	gtctttctgc	cctgctgtct	1620
	acatttttct	atggagcttg	tcataataat	agcagatatt	acctgatcag	gaatccctgt	1680
	ggcgcgtctg	acgctcatga	gtttttcatg	atggtgatga	gtagcactgc	actgtcacct	1740
	gatgattggc	cctgtctcgt	ttcccttctc	tctggggaga	tatgctgctt	ttccaccaga	1800
	cttgcctccat	actagaagct	tcttttgggt	tcaattaaaa	agaaaataag	ctagtcattc	1860
20	tgggcagcat	tttattgata	gaagggggaa	aaagtcattt	ctacttgcat	gatttttttaa	1920
	attaaattaa	attaaattaa	tttaaaaa				1948

&lt;210&gt; 4

&lt;211&gt; 1486

25 &lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 4

30	ccgcgcgcgc	gtccgggggc	tccagtaggg	ctgacgctcc	ggtgctcgca	caatcccccg	60
	cctcggcttg	caacggggcg	ccctccactc	cccaggtccc	cggcagccgc	cgccacccca	120
	gcgcgccccg	atctggcccc	ctgccccgcg	aagatggctg	ccgtacgccg	ggcccgcagt	180
	tattgccgct	gcctggtgag	cttctccgac	cgagaactct	gctaagctcc	gctgcagaga	240
	caggcaggag	tagacacccg	gacacccagc	acccctctct	cggggggcgg	tgacaggggg	300
	gcacggagag	cccctcgagc	gcagcaggcc	gccccgccag	catggcagaa	gctgaggaa	360
35	attgtcattc	tgatactgtc	agagcagatg	atgatgaaga	aaatgaaagt	cctgctgaaa	420
	cagatctgca	ggcacaactc	cagatgttcc	gagctcagtg	gatgtttgaa	cttgctccag	480
	gtgtaagctc	tagcaattta	gaaaatcgac	cttgacagag	agcaagaggc	tctctccaga	540
	aaacatcggc	agataccaaa	ggaaaacaa	aacaggcaaa	agaagaaaag	gctcgagaac	600
	tcttcttaaa	agcagtagaa	gaagaacaaa	atggagctct	ctatgaagcc	atcaagtttt	660
40	atcgtagggc	tatgcaactt	gtacctgata	tagagttcaa	gattacttat	accggtctc	720
	cagatgggtg	tggcgttggg	aacagctaca	ttgaagataa	tgatgatgac	agcaaaatgg	780
	cagatctctt	gtcctacttc	cagcagcaac	tcacatttca	ggagtctgtg	cttaactctg	840
	gtcagcctga	gcttgagagc	agtacagattc	acatatcagt	gctgccaatg	gaggtcctga	900
	gtacatctt	ccgatgggtg	gtgtctagt	acttggacct	cagatcattg	gagcagttgt	960
45	cgctggtgtg	cagaggattc	tacatctgtg	ccagagacc	tgaaatatgg	cgctcggcct	1020
	gcttgaaagt	ttggggcaga	agctgtatta	aacttgttcc	gtacacgtcc	tggagagaga	1080
	tgtttttaga	acggcctcgt	gttcgggttg	atggcggtga	tatcagtaaa	accacatata	1140
	ttcgtcaagg	ggaacagtct	cttgatgggt	tctatagagc	ctggcaccaa	gtggaatatt	1200
	acaggtagat	aagattcttt	cctgatggcc	atgtgatgat	gttgacaacc	cctgaagagc	1260
50	ctcagtccat	tgttccacgt	ttagaacta	gggaatacca	ggactgatgc	caattctact	1320
	gggtcactat	cgcttggcac	aagacacaga	ccatcagacc	caaagtattt	tgctgtaata	1380
	actaagaaaa	aaaggaagaa	aaaccacttg	gactataaat	acccgatatt	ttcgtcgggg	1440
	tccctgtacc	aaagaagccg	aatcagaagt	tttcatgnng	gggctn		1486

55 &lt;210&gt; 5

&lt;211&gt; 774

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

60 &lt;400&gt; 5

65	tgccgccccc	cacccccc	caccccccgc	atgcaacggg	attgaagggt	cctgccgggtg	60
	ggaccctgtc	cggcccgatg	ccactgcccc	ccgaggtctg	tagacgtagg	tgtaggcac	120
	gtcccccaca	ccgcgcgcct	cccacggcac	ctcggggaca	ccagagctgc	cgacttggag	180
	actcctggtc	tgtagagagc	cgggtggtgc	cgtgcccgca	ggaactgggc	tgggcctcgt	240
	gcgcccgtgg	ggtctgcgct	tggtctttct	gtgcttggat	ttgcataatt	attgcattgc	300
	tggtagagac	ccccaggcct	gtccaccctg	ccaagactcc	tcaggcagcg	tgtgggtccc	360
	gcactctgcc	ccattttccc	cgatgtcccc	tgcgggcgcg	ggcagccacc	caagcctgct	420

	ggctgcggcc	ccctctcggc	caggcattgg	ctcagcccg	tgagtggggg	gtcgtggggc	480
	agtccccgag	gagctggggc	cctgcacagg	cacacagggc	ccggccacac	ccagcggccc	540
	cccgccacagc	cacccgtggg	gtgctgcctt	tatgcccggc	gcccggcacc	aactccatgt	600
	ttggtgtttg	tctgtgtttg	ttttccaaga	aatgattcaa	attgctgctt	ggattttgaa	660
5	atttactgta	actgtcagtg	tacacgtctg	gaccccggtt	cattttttaca	ccaattttgg	720
	aaaaatgctg	ctctcagcct	cccacaatta	aaccgcatgt	gatctccaaa	aaaa	774

&lt;210&gt; 6

&lt;211&gt; 2936

10

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 6

	cgtaaccctt	agtcccaatg	cctccgtaag	cggagttgag	tgggtgcctg	tggttgagac	60
15	tgtggaggtg	tccccggtgg	cgagcgcggc	cagaactgcg	gtcacttaag	ttttccgtgt	120
	gcgggttgca	aggagcgtgc	gtgcgtctgg	tataggaggg	acacctctgg	attgaggatc	180
	ttatgacctt	cttttagagga	aggtataatt	tggcttctctg	agattctgcc	ttagcaagaa	240
	aggagtggga	aatacccttg	gaaagaaaac	taaaacagta	agaaagccaa	aactttatgt	300
	tacatggttg	tcagcacatt	taccgatatg	gacacttttc	ccaataattt	tcctcctggt	360
20	ggagacagtg	gattgacagg	ttctcagtcg	gagttccaga	aaatgttaat	tgatgaaagg	420
	ttacgatgtg	agcatcataa	agctaattat	cagacactga	aggctgaaca	cacaagggtg	480
	cagaatgaa	atgtaaagt	acaaaatgaa	ctcaagcacc	tgtttaatga	aaagcaaaact	540
	cagcaggaaa	aacttcagct	cctgcttgaa	gaactaagag	gagaattagt	agagaaaact	600
	aaagatttag	aagaaatgaa	actgcagata	ttaactccca	aaaattggaa	ttgctaagag	660
25	cccaaataca	acaagaatta	gaaactccaa	tgagagaacg	ttttaggaat	ctagatgaag	720
	aggtagaaaa	gtatagagct	gtatataata	agcttcgcta	tgaacataca	tttctcaagt	780
	cagaatttga	acaccagaag	gaagagtatg	cacgtatttt	agatgaagga	aaaataaaat	840
	atgaatcaga	gattgcaaga	ctggaggaag	ataaagaaga	actacgtaac	cagctgctta	900
	atgttgatct	cacaaaagac	agcaaacgag	tggaacaact	tgctcgagaa	aaagtctatt	960
30	tgtgtcaaaa	attaaaagg	ttagaggctg	aagtagcgga	attaaaggct	gaaaaggaga	1020
	attctgaggc	tcagggtgaa	aatgcccaaa	gaatacagg	gcggcagttg	gctgagatgc	1080
	aggctacagt	cagatccctg	ggggctgaaa	aacaatcagc	taattttacg	gcagaacgct	1140
	tggaaaaaaga	gctacaatca	agcagtgaac	aaaatacctt	tttaattaat	aaattgcata	1200
	aagctgaacg	agaaataaat	acattgtcca	gtaaagtaaa	agaacttaaa	cattcaaaac	1260
35	aactggaaat	aacagacatc	aaactggaga	cagcaagagc	taagagtga	ctagaaaag	1320
	aaaggaataa	gcttcaaaat	gaactggatg	gattacagtc	agacaatgaa	attctcaaa	1380
	cagctgttga	acatcacaaa	gtgctcttag	tagaaaagga	tcgtgaatta	atacgtaaa	1440
	tacaagctgc	caaagaagaa	ggttatcaaa	aacttgtggt	attacaagat	gaaaagttag	1500
	aactcgagaa	cgatttagca	gatttggaag	aaatgaaagt	ggaacatgat	gtctggaggc	1560
40	aatctgaaaa	ggatcagtat	gaagagaaat	tgccggcttc	acagatggca	gaagagatca	1620
	ccaggaagga	gcttcagagt	gttaggttaa	aacttcagca	acaaattgtg	actattgaaa	1680
	atgcagagaa	ggaaaaaaat	gaaaattctg	acctaataca	gcaaatcagt	agtttgca	1740
	tccaagtga	ttcacttgca	cagtcagaga	atgacttgct	gaattcaaac	caaattgctga	1800
	aggaataggt	ggagagatta	aaacaagaat	accgaaattt	tagaagccaa	gctgaaaaag	1860
45	cgcaactaga	agctgaaaag	acattggaag	agaaaacag	acagtgggtg	gaagaaaagc	1920
	ataagcttca	tgaccgtatc	acagacagag	aagaaaagta	caatcaagct	aaggagaaac	1980
	tgacgcgagc	tgcaattgcc	cagaaaaaga	gaaaatctct	tcatgaaaac	aaattgaaaa	2040
	gactacaaga	gaaagtagaa	gtcttgagg	caaagaaaga	agaattggaa	acagaaatca	2100
	ggctcttaaat	agacaaaaatg	ttcctttgaa	gactatacaa	ggcttcaaaa	aagactaaaa	2160
50	gatatacaga	gaagacataa	tgaatttcga	agtctaattt	tggttcctaa	catgcctcca	2220
	acagcatcta	tcaatcctgt	tagctttcag	tcacagacca	tggttcctaa	catggaacta	2280
	ccatttcttc	ctcatatgca	ggaggaacaa	catcaaagg	aactctctct	acttcgcaaa	2340
	agactagaag	aactggaac	aacacaaaga	aaacaactag	aggaacttgg	atcttccgga	2400
	gaatgatgtt	cttgagagac	aggcagatca	aaagaggtga	agttgggtgac	tcagtaaaac	2460
55	ttgacatttt	aacctgtggc	atttagatac	tttttactgt	ttgccaaaac	acttgaatgt	2520
	gcctcaagaa	aaggtaccta	ctacatgctg	tattgtatga	ctgtcaggat	tttaagatta	2580
	tacaagtga	gcattaaaa	agaaattctc	agagataatt	agaatatttg	acaatgggtt	2640
	gagaatgtaa	aacaaaaagg	aactagttag	agtcaagttt	taaaattttt	actttgttga	2700
	atTTTTTTTT	ttggcatttt	gagtgaataa	taactatcat	taattctctc	ttcatctttg	2760
60	agatgcttgg	ccataacagg	gtccatacac	atcttctggt	ttactatata	caaaaactgt	2820
	agttgaaaaa	agatgacaat	ttaaaagtca	gcctaaagaa	tgtaaaggta	tctatataca	2880
	aaaggctacc	ttttctaaaa	tctgtgtgca	cataattaaa	gagcttaatt	tttaaa	2936

&lt;210&gt; 7

65

&lt;211&gt; 1387

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

<400> 7

	aaaggctaca	tcattaacac	tagaaggagg	acgattaaaa	cgaactccac	agctgattca	60
	tggaagagac	tatgaaatgg	tcccagaacc	tgtgtggaga	gcactttatc	actggtatgg	120
5	agcaaacctg	gccttaccta	gaccagttat	caagaacagc	aagacagaca	tcccagagct	180
	ggaattatct	ccccgctatc	ttctcttctc	gagacagcag	cctgccactc	ggacacagca	240
	gtctaacatc	tgggtgaata	tgggaaatgt	accttctccg	aatgcacctt	taaagcgggt	300
	attagcctat	acaggctggt	ttagtgcgaat	gcagaccatc	aaggaaaatc	acgaatatct	360
	atctcaaagg	ctgcgcatca	aagaggaaga	tatgcgctcg	tggctatata	acagtgcgaa	420
10	ctaccttact	cttctggatg	atgaggatca	taaattggaa	tatttgaaaa	tccaggatga	480
	acaacacctg	gtaattgaag	ttcgcaacaa	agatatgagt	tggcctgagg	agatgtcttt	540
	tatagcaaat	agtagtaaaa	tagatagaca	caaggttccc	acagaaaagg	gagccacagg	600
	tctaagcaat	ctgggaaaca	catgcttcat	gaactcaagc	atccagtgtg	ttagtaacac	660
	acagccactg	acacagtatt	ttatctcagg	tgagcatctt	tatgaactca	acaggacaaa	720
15	tcccattggt	atgaaggggc	atatggctaa	atgctatggt	gatttagtgc	aggaactttg	780
	gagtggaaat	cagaagaatg	ttgccccatt	aaagcttcgg	tggaccatag	caaaatatgc	840
	tcccagggtt	aatgggtttc	agcaacagga	ctcccaagaa	cttctggttt	ttctcttgga	900
	tggtcttcat	gaagatctta	atcgagtcca	tgaaaagcca	tatgtggaac	tgaaggacag	960
	tgatgggcga	ccagactggg	aagtagctgc	agagccctgg	gacaaccatc	taagaagaaa	1020
20	tagatcaatt	gttgtggatt	gtttccatgg	gcagctaaga	tctcaagtaa	aatgcaagac	1080
	atgtgggcat	ataagtgtcc	gatttgaccc	tttcaatttt	ttgtctttgc	caactaccaat	1140
	ggacagttat	atgcacttag	aaataacagt	gattaagtta	gatggtacta	cccctgtacg	1200
	gtatggacta	agactgaata	tggatgaaaa	gtacacagggt	ttaaaaaaac	agctgagtga	1260
	tctctgtgga	cttaattcag	aacaaatcct	tctagcagaa	gtacatgggt	ccaacataaa	1320
25	gaactttcct	caggacaacc	caaaaagtac	cgaacttctc	agtgaagtgg	gatttttttg	1380
	tgtgccca						1387

<210> 8  
 <211> 749  
 <212> DNA  
 <213> Homo sapiens

<400> 8

	gccaacatg	gctggagcgc	ggcggagggtg	agccggccgc	ccgccgcgag	acgccccagc	60
35	ctactgcgcc	cgagtcgccg	ggccccagtg	gcgcctcagc	tctgcgggtg	cgaggcccaa	120
	cggctcgatc	gctgcccgcc	gccagcatgt	tgggcgcccc	ggacgagagc	tccgtgcggg	180
	tggctgtcag	aataagacca	cagcttgcca	aagagaagat	tgaaggatgc	catatttgta	240
	catctgtcac	accaggagag	cctcaggtct	tcctagggaa	agataaggct	tttacttttg	300
	actatgtatt	tgacattgac	tcccagcaag	agcagatcta	cattcaatgt	atagaaaaac	360
40	taattgaagg	ttgctttgaa	ggatacaatg	ctacagtttt	tgcttatgga	caaactggag	420
	ctggtaaaac	atacacaatg	ggaacaggat	ttgatgttaa	cattgttgag	gaagaactgg	480
	gtattatttc	tcgagctgtt	aaacaccttt	ttaagagtat	tgaagaaaaa	aaacacatag	540
	caattaaaaa	tgggcttctc	gctccagatt	ttaaagtga	tgcccaattc	ttagagctct	600
	ataatgaaga	ggtccttgac	ttatttgata	ccactcgtga	tattgatgca	aaaagtaaaa	660
45	aatcaaatat	aagaattcat	gaagattcaa	ctggagggaa	tttatactgt	gggcgtttcc	720
	aacacgtact	gtgaatacag	aatcnagag				749

<210> 9  
 <211> 686  
 <212> DNA  
 <213> Homo sapiens

<400> 9

	tggaaattat	agacctagca	aaaaaagatt	tagagaaggt	tgnaaagaaa	agaaaagagg	60
55	aagnaaaaaa	gtgtggctgg	taaagaggat	aatacagaca	ctgaccaaga	gaagaaaagaa	120
	gaaaagggtg	tttcggaaag	agaaacccaa	tgaattagaa	gtggaagaaa	gtcaagaagt	180
	gagtgatcat	gaggatgaag	aagaggagga	ggaggaggag	gaagatgaca	ttgatggggg	240
	tgaagttct	gatgaatcag	attctgaatc	agatgaaaaa	gccnattatc	aagcagactt	300
	ggcaaacatt	acttgtgaaa	ttgcaattaa	gcaaaagctg	attgatgaac	tagaaaaacg	360
60	ccagaaaaga	ctgcagactc	tgaaaaagca	gtatgaagag	aagntaatga	tgctgcaaca	420
	taaaattcgg	gatactcagc	ttgaaaagaga	ccagggtgct	caaaacttag	gctcggtaga	480
	atcttactca	gaagaaaaag	caaaaaaagt	taggtctgaa	tatgaaaaga	aactccaagc	540
	catgaacaaa	gaactgcaga	gacttcaagc	agctcaaaaa	gaacatgcaa	ggttgcttaa	600
	aaatcagtct	cagtatgaaa	agcnattgaa	gaaattgcag	caggatgtga	tggaaatgaa	660
65	aaaaacaaa	gttcgcctaa	tgaaaa				686

<210> 10



<211> 833  
 <212> DNA  
 <213> Homo sapiens

5       <400> 10  
 gttcttctgt cgccggcttc agcagcccg ccccgggcag gaatagaaga tgaacaaacc 60  
 cataacacca tcaacatatg tgcgctgcct caatgttgga ctaattagga agctgtcaga 120  
 ttttattgat cctcaagaag gatggaagaa gttagctgta gctattaaaa aaccatctgg 180  
 10   tgaatgata tacaatcagt ttcacataag gagatttgaa gcattacttc aaactggaaa 240  
 aagtcccact tctgaattac tgtttgactg gggcaccaca aattgcacag ttgggtgatct 300  
 tgtggatctt ttgatccaaa atgaattttt tgctcctgcg agtcttttgc tcccagatgc 360  
 tgttcccaaa actgctaata cactaccttc taaagaagct ataacagttc agcaaaaaa 420  
 gatgcctttc tgtgacaaag acaggacatt gatgacacct gtgcagaatc ttgaacaaag 480  
 ctatatgccca cctgactcct caagtccaga aaataaaaagt ttagaagtta gtgatacacg 540  
 15   ttttcacagt ttttcatttt atgaattgaa gaatgtcaca aataactttg atgaacgacc 600  
 catttctgtt ggtggtaata aaatgggaga gggaggattt ggagttgtat ataaagggct 660  
 acgtaaataa cacaactgtg gcagtgaag aagcttgacg caatgggtga cttactact 720  
 gaaggaactg aaaccagcag tttgatccaa gaaaataaaa gtaatgggca aaagtggta 780  
 accatggaaa aacttttagta ggaacctact tgggttttct caagtggatg gga 833

<210> 11  
 <211> 799  
 <212> DNA  
 <213> Homo sapiens

25       <400> 11  
 taaaaatc cccttngatg atacctgccataatgatgt gtccattat tagattatgt 60  
 tacatgccaa agtttaaaagg aatttgggca gatccagtta aggttcctta aacaacntca 120  
 ctttgagact gttgaaagg cctgacctaa tccaagtga ccccttgcaa gaagaattct 180  
 30   ccttgtaagc cttgaagaag tatgtgagag ggccacattg gctaaaaact aaaggtggcc 240  
 tctaggagat gagacctacc ttccagttgt cagcaagcag gaaaaaaaaa ttgggacctc 300  
 agttgcaacc acaaggaaact gaattctgcc aaaaatntga gtcagcttag aagagtactc 360  
 caagcttcag atgataacca cagcctgggc tgacacctgg atttagcttt gcatgatcct 420  
 cagtatgaga atctatctgt tctgtgctgg acttctaata tatagaactg tgagataatg 480  
 35   ggtcacattg gctggatgtg gtggctcata cctgtaaatc ccagcacttt gggaggccga 540  
 ggcaggcaga tcacctgagg tcaagagttc aagaccggcc tggccaacat ggtgaaaccc 600  
 cgtctntact aaaaatacaa aaattagacg agcgtgggtg tggacacctg tagtccagc 660  
 tgcttgggag gctgaggcag gagactagct ggaaccaggg aggtagaggt tgcagtgagc 720  
 tgagatcgtg cactgcact ccagcctggg tgacagagtg agactccatc ataaataaat 780  
 40   aaataaataa atgggtcnc 799

<210> 12  
 <211> 812  
 <212> PRT  
 <213> Homo sapiens

45       <400> 12  
 Met Ala Gly Gly His Cys Gly Ser Phe Pro Ala Ala Ala Ala Gly Ser  
 1                   5                   10                   15  
 50   Gly Glu Ile Val Gln Leu Asn Val Gly Gly Thr Arg Phe Ser Thr Ser  
                  20                   25                   30  
 Arg Gln Thr Leu Met Trp Ile Pro Asp Ser Phe Phe Ser Ser Leu Leu  
                  35                   40                   45  
 55   Ser Gly Arg Ile Ser Thr Leu Arg Asp Glu Thr Gly Ala Ile Phe Ile  
                  50                   55                   60  
 Asp Arg Asp Pro Ala Ala Phe Ala Pro Ile Leu Asn Phe Leu Arg Thr  
 65                   70                   75                   80  
 Lys Glu Leu Asp Leu Arg Gly Val Ser Ile Asn Val Leu Arg His Glu  
                  85                   90                   95  
 60   Ala Glu Phe Tyr Gly Ile Thr Pro Leu Val Arg Arg Leu Leu Leu Cys  
                  100                   105                   110  
 Glu Glu Leu Glu Arg Ser Ser Cys Gly Ser Val Leu Phe His Gly Tyr  
                  115                   120                   125  
 65   Leu Pro Pro Pro Gly Ile Pro Ser Arg Lys Ile Asn Asn Thr Val Arg  
                  130                   135                   140  
 Ser Ala Asp Ser Arg Asn Gly Leu Asn Ser Thr Glu Gly Glu Ala Arg  
                  145                   150                   155                   160

Gly Asn Gly Thr Gln Pro Val Leu Ser Gly Thr Gly Glu Glu Thr Val  
 165 170 175  
 Arg Leu Gly Phe Pro Val Asp Pro Arg Lys Val Leu Ile Val Ala Gly  
 180 185 190  
 5 His His Asn Trp Ile Val Ala Ala Tyr Ala His Phe Ala Tyr Arg Ile  
 195 200 205  
 Lys Glu Ser Ser Gly Trp Gln Gln Val Phe Thr Ser Pro Tyr Leu Asp  
 210 215 220  
 10 Trp Thr Ile Glu Arg Val Ala Leu Asn Ala Lys Val Val Gly Gly Pro  
 225 230 235 240  
 His Gly Asp Lys Asp Lys Met Val Ala Val Ala Ser Glu Ser Ser Ile  
 245 250 255  
 Ile Leu Trp Ser Val Gln Asp Gly Gly Ser Gly Ser Glu Ile Gly Val  
 260 265 270  
 15 Phe Ser Leu Gly Val Pro Val Asp Ala Leu Phe Phe Ile Gly Asn Gln  
 275 280 285  
 Leu Val Ala Thr Ser His Thr Gly Lys Val Gly Val Trp Asn Ala Val  
 290 295 300  
 20 Thr Gln His Trp Gln Val Gln Asp Val Val Pro Ile Thr Ser Tyr Asp  
 305 310 315 320  
 Thr Ala Gly Ser Phe Leu Leu Leu Gly Cys Asn Asn Gly Ser Ile Tyr  
 325 330 335  
 Tyr Ile Asp Met Gln Lys Phe Pro Leu Arg Met Lys Asp Asn Asp Leu  
 340 345 350  
 25 Leu Val Thr Glu Leu Tyr His Asp Pro Ser Asn Asp Ala Ile Thr Ala  
 355 360 365  
 Leu Ser Val Tyr Leu Thr Pro Lys Thr Ser Val Ser Gly Asn Trp Ile  
 370 375 380  
 30 Glu Ile Ala Tyr Gly Thr Ser Ser Gly Ala Val Arg Val Ile Val Gln  
 385 390 395 400  
 His Pro Glu Thr Val Gly Ser Gly Pro Gln Leu Phe Gln Thr Phe Thr  
 405 410 415  
 Val His Arg Ser Pro Val Thr Lys Ile Met Leu Ser Glu Lys His Leu  
 420 425 430  
 35 Val Ser Val Cys Ala Asp Asn Asn His Val Arg Thr Trp Thr Val Thr  
 435 440 445  
 Arg Phe Gly Met Ile Ser Thr Gln Pro Gly Ser Thr Pro Leu Ala Ser  
 450 455 460  
 40 Phe Lys Ile Leu Ser Leu Glu Glu Thr Glu Ser His Gly Ser Tyr Ser  
 465 470 475 480  
 Ser Gly Asn Asp Ile Gly Pro Phe Gly Glu Arg Asp Asp Gln Gln Val  
 485 490 495  
 Phe Ile Gln Lys Val Val Pro Ile Thr Asn Lys Leu Phe Val Arg Leu  
 500 505 510  
 45 Ser Ser Thr Gly Lys Arg Ile Cys Glu Ile Gln Ala Val Asp Cys Thr  
 515 520 525  
 Thr Ile Ser Ser Phe Thr Val Arg Glu Cys Glu Gly Ser Ser Arg Met  
 530 535 540  
 50 Gly Ser Arg Pro Arg Arg Tyr Leu Phe Thr Gly His Thr Asn Gly Ser  
 545 550 555 560  
 Ile Gln Met Trp Asp Leu Thr Thr Ala Met Asp Met Val Asn Lys Ser  
 565 570 575  
 Glu Asp Lys Asp Val Gly Gly Pro Thr Glu Glu Glu Leu Leu Lys Leu  
 580 585 590  
 55 Leu Asp Gln Cys Asp Leu Ser Thr Ser Arg Cys Ala Thr Pro Asn Ile  
 595 600 605  
 Ser Pro Ala Thr Ser Val Val Gln His Ser His Leu Arg Glu Ser Asn  
 610 615 620  
 60 Ser Ser Leu Gln Leu Gln His His Asp Thr Thr His Glu Ala Ala Thr  
 625 630 635 640  
 Tyr Gly Ser Met Arg Pro Tyr Arg Glu Ser Pro Leu Leu Ala Arg Ala  
 645 650 655  
 Arg Arg Thr Glu Ser Phe His Ser Tyr Arg Asp Phe Gln Thr Ile Asn  
 660 665 670  
 65 Leu Asn Arg Asn Val Glu Arg Ala Val Pro Glu Asn Gly Asn Leu Gly  
 675 680 685  
 Pro Ile Gln Ala Glu Val Lys Gly Ala Thr Gly Glu Cys Asn Ile Ser

690 695 700  
 Glu Arg Lys Ser Pro Gly Val Glu Ile Lys Ser Leu Arg Glu Leu Asp  
 705 710 715 720  
 Ser Gly Leu Glu Val His Lys Ile Ala Glu Gly Phe Ser Glu Ser Lys  
 5 725 730 735  
 Lys Arg Ser Ser Glu Asp Glu Asn Glu Asn Lys Ile Glu Phe Arg Lys  
 740 745 750  
 Lys Gly Gly Phe Glu Gly Gly Gly Phe Leu Gly Arg Lys Lys Val Pro  
 755 760 765  
 10 Tyr Leu Ala Ser Ser Pro Ser Thr Ser Asp Gly Gly Thr Asp Ser Pro  
 770 775 780  
 Gly Thr Ala Ser Pro Ser Pro Thr Lys Thr Thr Pro Ser Pro Arg His  
 785 790 795 800  
 15 Lys Lys Ser Asp Ser Ser Gly Gln Glu Tyr Ser Leu  
 805 810

<210> 13  
 <211> 167  
 <212> PRT  
 20 <213> Homo sapiens

<400> 13  
 Leu Ala Ser Ala Arg Ile Pro Gly Arg Pro Tyr Val Gly Val Leu Arg  
 1 5 10 15  
 25 Pro Pro Gly Ser Leu Phe Phe Ala Pro Ala Gly Thr Thr Ile Ser Ala  
 20 25 30  
 Leu Leu Gly Phe Cys Tyr Met Ala Ala Phe Leu Lys Met Ser Val Ser  
 35 40 45  
 Val Asn Phe Phe Arg Pro Phe Thr Arg Phe Leu Val Pro Phe Thr Leu  
 50 55 60  
 30 His Arg Lys Arg Asn Asn Leu Thr Ile Leu Gln Arg Tyr Met Ser Ser  
 65 70 75 80  
 Lys Ile Pro Ala Val Thr Tyr Pro Lys Asn Glu Ser Thr Arg Pro Ser  
 85 90 95  
 35 Glu Glu Leu Glu Leu Asp Lys Trp Lys Thr Thr Met Lys Ser Val  
 100 105 110  
 Gln Glu Glu Cys Val Ser Thr Ile Ser Ser Ser Lys Asp Glu Asp Pro  
 115 120 125  
 40 Leu Ala Ala Thr Arg Glu Phe Ile Glu Met Trp Arg Leu Leu Gly Arg  
 130 135 140  
 Glu Val Pro Glu His Ile Thr Glu Glu Glu Leu Lys Thr Leu Met Glu  
 145 150 155 160  
 Cys Val Ser Asn Thr Ala Lys  
 165

<210> 14  
 <211> 452  
 <212> PRT  
 45 <213> Homo sapiens

50 <400> 14  
 Pro Cys Glu Gly Gly Ala Arg Ser Cys Leu Val Thr Glu Ser Ala Arg  
 1 5 10 15  
 55 Gly Gly Leu Gln Phe Leu Gln Gln Cys Asp Arg Glu Asp Leu Val Glu  
 20 25 30  
 Leu Ala Leu Pro Gln Leu Ala Gln Val Val Thr Val Tyr Glu Phe Leu  
 35 40 45  
 Leu Met Lys Val Glu Lys Asp His Leu Ala Lys Pro Phe Phe Pro Ala  
 50 55 60  
 60 Ile Tyr Lys Glu Phe Glu Glu Leu His Lys Met Val Lys Lys Met Cys  
 65 70 75 80  
 Gln Asp Tyr Leu Ser Ser Ser Gly Leu Cys Ser Gln Glu Thr Leu Glu  
 85 90 95  
 65 Ile Asn Asn Asp Lys Val Ala Glu Ser Leu Gly Ile Thr Glu Phe Leu  
 100 105 110  
 Arg Lys Lys Glu Ile His Pro Asp Asn Leu Gly Pro Lys His Leu Ser  
 115 120 125

Arg Asp Met Asp Gly Glu Gln Leu Glu Gly Ala Ser Ser Glu Lys Arg  
 130 135 140  
 Glu Arg Glu Ala Ala Glu Glu Gly Leu Ala Ser Val Lys Arg Pro Arg  
 145 150 155 160  
 5 Arg Glu Ala Leu Ser Asn Asp Thr Thr Glu Ser Leu Ala Ala Asn Ser  
 165 170 175  
 Arg Gly Arg Glu Lys Pro Arg Pro Leu His Ala Leu Pro Ala Gly Phe  
 180 185 190  
 10 Ser Pro Pro Val Asn Val Thr Val Ser Pro Arg Ser Glu Glu Ser His  
 195 200 205  
 Thr Thr Thr Val Ser Gly Gly Asn Gly Ser Val Phe Gln Ala Gly Pro  
 210 215 220  
 Gln Leu Gln Ala Leu Ala Asn Leu Glu Ala Arg Arg Gly Ser Ile Gly  
 225 230 235 240  
 15 Ala Ala Leu Ser Ser Arg Asp Val Ser Gly Leu Pro Val Tyr Ala Gln  
 245 250 255  
 Ser Gly Glu Pro Arg Arg Leu Thr Gln Ala Gln Val Ala Ala Phe Pro  
 260 265 270  
 20 Gly Glu Asn Ala Leu Glu His Ser Ser Asp Gln Asp Thr Trp Asp Ser  
 275 280 285  
 Leu Arg Ser Pro Gly Phe Cys Ser Pro Leu Ser Ser Gly Gly Gly Ala  
 290 295 300  
 Glu Ser Leu Pro Pro Gly Gly Pro Gly His Ala Glu Ala Gly His Leu  
 305 310 315 320  
 25 Gly Lys Val Cys Asp Phe His Leu Asn His Gln Gln Pro Ser Pro Thr  
 325 330 335  
 Ser Val Leu Pro Thr Glu Val Ala Ala Pro Pro Leu Glu Lys Ile Leu  
 340 345 350  
 30 Ser Val Asp Ser Val Ala Val Asp Cys Ala Tyr Arg Thr Val Pro Lys  
 355 360 365  
 Pro Gly Pro Gln Pro Gly Pro His Gly Ser Leu Leu Thr Glu Gly Cys  
 370 375 380  
 Leu Arg Ser Leu Ser Gly Asp Leu Asn Arg Phe Pro Cys Gly Met Glu  
 385 390 395 400  
 35 Val His Ser Gly Gln Arg Glu Leu Glu Ser Val Val Ala Val Gly Glu  
 405 410 415  
 Ala Met Ala Leu Lys Phe Pro Met Gly Ala Met Ser Tyr Cys Leu Arg  
 420 425 430  
 40 Asp Arg Ser Arg Phe Leu Phe Arg Leu Pro Met Gly Leu Ser Cys Pro  
 435 440 445  
 Leu Gln Val Gln  
 450  
 <210> 15  
 <211> 321  
 <212> PRT  
 <213> Homo sapiens  
 <400> 15  
 50 Met Ala Glu Ala Glu Glu Asp Cys His Ser Asp Thr Val Arg Ala Asp  
 1 5 10 15  
 Asp Asp Glu Glu Asn Glu Ser Pro Ala Glu Thr Asp Leu Gln Ala Gln  
 20 25 30  
 55 Leu Gln Met Phe Arg Ala Gln Trp Met Phe Glu Leu Ala Pro Gly Val  
 35 40 45  
 Ser Ser Ser Asn Leu Glu Asn Arg Pro Cys Arg Ala Ala Arg Gly Ser  
 50 55 60  
 Leu Gln Lys Thr Ser Ala Asp Thr Lys Gly Lys Gln Glu Gln Ala Lys  
 65 70 75 80  
 60 Glu Glu Lys Ala Arg Glu Leu Phe Leu Lys Ala Val Glu Glu Glu Gln  
 85 90 95  
 Asn Gly Ala Leu Tyr Glu Ala Ile Lys Phe Tyr Arg Arg Ala Met Gln  
 100 105 110  
 65 Leu Val Pro Asp Ile Glu Phe Lys Ile Thr Tyr Thr Arg Ser Pro Asp  
 115 120 125  
 Gly Asp Gly Val Gly Asn Ser Tyr Ile Glu Asp Asn Asp Asp Asp Ser  
 130 135 140

Lys Met Ala Asp Leu Leu Ser Tyr Phe Gln Gln Gln Leu Thr Phe Gln  
 145 150 155 160  
 Glu Ser Val Leu Lys Leu Cys Gln Pro Glu Leu Glu Ser Ser Gln Ile  
 165 170 175  
 5 His Ile Ser Val Leu Pro Met Glu Val Leu Met Tyr Ile Phe Arg Trp  
 180 185 190  
 Val Val Ser Ser Asp Leu Asp Leu Arg Ser Leu Glu Gln Leu Ser Leu  
 195 200 205  
 Val Cys Arg Gly Phe Tyr Ile Cys Ala Arg Asp Pro Glu Ile Trp Arg  
 210 215 220  
 10 Leu Ala Cys Leu Lys Val Trp Gly Arg Ser Cys Ile Lys Leu Val Pro  
 225 230 235 240  
 Tyr Thr Ser Trp Arg Glu Met Phe Leu Glu Arg Pro Arg Val Arg Phe  
 245 250 255  
 15 Asp Gly Val Tyr Ile Ser Lys Thr Thr Tyr Ile Arg Gln Gly Glu Gln  
 260 265 270  
 Ser Leu Asp Gly Phe Tyr Arg Ala Trp His Gln Val Glu Tyr Tyr Arg  
 275 280 285  
 Tyr Ile Arg Phe Phe Pro Asp Gly His Val Met Met Leu Thr Thr Pro  
 290 295 300  
 20 Glu Glu Pro Gln Ser Ile Val Pro Arg Leu Arg Thr Arg Glu Tyr Gln  
 305 310 315 320  
 Asp

25  
 <210> 16  
 <211> 172  
 <212> PRT  
 <213> Homo sapiens

30  
 <400> 16  
 Ala Cys Pro Thr His Pro Pro Pro Pro Thr Ala Pro Arg Gly His Gln  
 1 5 10 15  
 35 Ser Cys Arg Leu Gly Asp Ser Trp Ser Val Lys Ser Arg Trp Cys Pro  
 20 25 30  
 Cys Pro Gln Glu Leu Gly Trp Ala Ser Cys Ala Arg Gly Val Cys Ala  
 35 40 45  
 Trp Ser Phe Cys Ala Trp Ile Cys Ile Phe Ile Ala Leu Leu Val Glu  
 50 55 60  
 40 Thr Pro Arg Pro Val His Pro Ala Lys Thr Pro Gln Ala Ala Cys Gly  
 65 70 75 80  
 Ser Arg Thr Leu Pro Pro Phe Pro Arg Cys Pro Leu Arg Ala Arg Ala  
 85 90 95  
 45 Ala Thr Gln Ala Cys Trp Leu Arg Pro Pro Leu Gly Gln Ala Leu Ala  
 100 105 110  
 Gln Pro Ala Glu Trp Gly Val Val Gly Gln Ser Pro Arg Ser Trp Ala  
 115 120 125  
 Pro Ala Gln Ala His Arg Ala Arg Pro His Pro Ala Ala Pro Arg Thr  
 130 135 140  
 50 Ala Thr Arg Gly Val Leu Pro Leu Cys Pro Ala Pro Gly Thr Asn Ser  
 145 150 155 160  
 Met Phe Gly Val Cys Leu Cys Leu Phe Phe Lys Lys  
 165 170

55  
 <210> 17  
 <211> 472  
 <212> PRT  
 <213> Homo sapiens

60  
 <400> 17  
 Met Arg Glu Arg Phe Arg Asn Leu Asp Glu Glu Val Glu Lys Tyr Arg  
 1 5 10 15  
 Ala Val Tyr Asn Lys Leu Arg Tyr Glu His Thr Phe Leu Lys Ser Glu  
 20 25 30  
 65 Phe Glu His Gln Lys Glu Glu Tyr Ala Arg Ile Leu Asp Glu Gly Lys  
 35 40 45  
 Ile Lys Tyr Glu Ser Glu Ile Ala Arg Leu Glu Glu Asp Lys Glu Glu

50 55 60  
 Leu Arg Asn Gln Leu Leu Asn Val Asp Leu Thr Lys Asp Ser Lys Arg  
 65 70 75 80  
 Val Glu Gln Leu Ala Arg Glu Lys Val Tyr Leu Cys Gln Lys Leu Lys  
 85 90 95  
 5 Gly Leu Glu Ala Glu Val Ala Glu Leu Lys Ala Glu Lys Glu Asn Ser  
 100 105 110  
 Glu Ala Gln Val Glu Asn Ala Gln Arg Ile Gln Val Arg Gln Leu Ala  
 115 120 125  
 10 Glu Met Gln Ala Thr Val Arg Ser Leu Gly Ala Glu Lys Gln Ser Ala  
 130 135 140  
 Asn Leu Arg Ala Glu Arg Leu Glu Lys Glu Leu Gln Ser Ser Ser Glu  
 145 150 155 160  
 15 Gln Asn Thr Phe Leu Ile Asn Lys Leu His Lys Ala Glu Arg Glu Ile  
 165 170 175  
 Asn Thr Leu Ser Ser Lys Val Lys Glu Leu Lys His Ser Asn Lys Leu  
 180 185 190  
 Glu Ile Thr Asp Ile Lys Leu Glu Thr Ala Arg Ala Lys Ser Glu Leu  
 195 200 205  
 20 Glu Arg Glu Arg Asn Lys Leu Gln Ser Glu Leu Asp Gly Leu Gln Ser  
 210 215 220  
 Asp Asn Glu Ile Leu Lys Ala Ala Val Glu His His Lys Val Leu Leu  
 225 230 235 240  
 25 Val Glu Lys Asp Arg Glu Leu Ile Arg Lys Val Gln Ala Ala Lys Glu  
 245 250 255  
 Glu Gly Tyr Gln Lys Leu Val Val Leu Gln Asp Glu Lys Leu Glu Leu  
 260 265 270  
 Glu Asn Arg Leu Ala Asp Leu Glu Lys Met Lys Val Glu His Asp Val  
 275 280 285  
 30 Trp Arg Gln Ser Glu Lys Asp Gln Tyr Glu Glu Lys Leu Arg Ala Ser  
 290 295 300  
 Gln Met Ala Glu Glu Ile Thr Arg Lys Glu Leu Gln Ser Val Arg Leu  
 305 310 315 320  
 35 Lys Leu Gln Gln Gln Ile Val Thr Ile Glu Asn Ala Glu Lys Glu Lys  
 325 330 335  
 Asn Glu Asn Ser Asp Leu Lys Gln Gln Ile Ser Ser Leu Gln Ile Gln  
 340 345 350  
 Val Thr Ser Leu Ala Gln Ser Glu Asn Asp Leu Leu Asn Ser Asn Gln  
 355 360 365  
 40 Met Leu Lys Glu Met Val Glu Arg Leu Lys Gln Glu Cys Arg Asn Phe  
 370 375 380  
 Arg Ser Gln Ala Glu Lys Ala Gln Leu Glu Ala Glu Lys Thr Leu Glu  
 385 390 395 400  
 45 Glu Lys Gln Ile Gln Trp Leu Glu Glu Lys His Lys Leu His Asp Arg  
 405 410 415  
 Ile Thr Asp Arg Glu Glu Lys Tyr Asn Gln Ala Lys Glu Lys Leu Gln  
 420 425 430  
 Arg Ala Ala Ile Ala Gln Lys Lys Arg Lys Ser Leu His Glu Asn Lys  
 435 440 445  
 50 Leu Lys Arg Leu Gln Glu Lys Val Glu Val Leu Glu Ala Lys Lys Glu  
 450 455 460  
 Glu Leu Glu Thr Glu Ile Arg Ser  
 465 470  
 55 <210> 18  
 <211> 462  
 <212> PRT  
 <213> Homo sapiens  
 60 <400> 18  
 Lys Ala Thr Ser Leu Thr Leu Glu Gly Gly Arg Leu Lys Arg Thr Pro  
 1 5 10 15  
 Gln Leu Ile His Gly Arg Asp Tyr Glu Met Val Pro Glu Pro Val Trp  
 20 25 30  
 65 Arg Ala Leu Tyr His Trp Tyr Gly Ala Asn Leu Ala Leu Pro Arg Pro  
 35 40 45  
 Val Ile Lys Asn Ser Lys Thr Asp Ile Pro Glu Leu Glu Leu Phe Pro

50 55 60  
 Arg Tyr Leu Leu Phe Leu Arg Gln Gln Pro Ala Thr Arg Thr Gln Gln  
 65 70 75 80  
 Ser Asn Ile Trp Val Asn Met Gly Asn Val Pro Ser Pro Asn Ala Pro  
 85 90 95  
 5 Leu Lys Arg Val Leu Ala Tyr Thr Gly Cys Phe Ser Arg Met Gln Thr  
 100 105 110  
 Ile Lys Glu Ile His Glu Tyr Leu Ser Gln Arg Leu Arg Ile Lys Glu  
 115 120 125  
 10 Glu Asp Met Arg Leu Trp Leu Tyr Asn Ser Glu Asn Tyr Leu Thr Leu  
 130 135 140  
 Leu Asp Asp Glu Asp His Lys Leu Glu Tyr Leu Lys Ile Gln Asp Glu  
 145 150 155 160  
 15 Gln His Leu Val Ile Glu Val Arg Asn Lys Asp Met Ser Trp Pro Glu  
 165 170 175  
 Glu Met Ser Phe Ile Ala Asn Ser Ser Lys Ile Asp Arg His Lys Val  
 180 185 190  
 Pro Thr Glu Lys Gly Ala Thr Gly Leu Ser Asn Leu Gly Asn Thr Cys  
 195 200 205  
 20 Phe Met Asn Ser Ser Ile Gln Cys Val Ser Asn Thr Gln Pro Leu Thr  
 210 215 220  
 Gln Tyr Phe Ile Ser Gly Arg His Leu Tyr Glu Leu Asn Arg Thr Asn  
 225 230 235 240  
 25 Pro Ile Gly Met Lys Gly His Met Ala Lys Cys Tyr Gly Asp Leu Val  
 245 250 255  
 Gln Glu Leu Trp Ser Gly Thr Gln Lys Asn Val Ala Pro Leu Lys Leu  
 260 265 270  
 Arg Trp Thr Ile Ala Lys Tyr Ala Pro Arg Phe Asn Gly Phe Gln Gln  
 275 280 285  
 30 Gln Asp Ser Gln Glu Leu Leu Ala Phe Leu Leu Asp Gly Leu His Glu  
 290 295 300  
 Asp Leu Asn Arg Val His Glu Lys Pro Tyr Val Glu Leu Lys Asp Ser  
 305 310 315 320  
 35 Asp Gly Arg Pro Asp Trp Glu Val Ala Ala Glu Ala Trp Asp Asn His  
 325 330 335  
 Leu Arg Arg Asn Arg Ser Ile Val Val Asp Leu Phe His Gly Gln Leu  
 340 345 350  
 Arg Ser Gln Val Lys Cys Lys Thr Cys Gly His Ile Ser Val Arg Phe  
 355 360 365  
 40 Asp Pro Phe Asn Phe Leu Ser Leu Pro Leu Pro Met Asp Ser Tyr Met  
 370 375 380  
 His Leu Glu Ile Thr Val Ile Lys Leu Asp Gly Thr Thr Pro Val Arg  
 385 390 395 400  
 45 Tyr Gly Leu Arg Leu Asn Met Asp Glu Lys Tyr Thr Gly Leu Lys Lys  
 405 410 415  
 Gln Leu Ser Asp Leu Cys Gly Leu Asn Ser Glu Gln Ile Leu Leu Ala  
 420 425 430  
 Glu Val His Gly Ser Asn Ile Lys Asn Phe Pro Gln Asp Asn Pro Lys  
 435 440 445  
 50 Ser Thr Glu Leu Leu Ser Glu Val Gly Phe Phe Gly Val Pro  
 450 455 460  
  
 <210> 19  
 <211> 243  
 55 <212> PRT  
 <213> Homo sapiens  
  
 <400> 19  
 60 Pro Thr Trp Leu Glu Arg Gly Gly Gly Glu Pro Ala Ala Arg Pro Gln  
 1 5 10 15  
 Thr Pro Gln Pro Thr Ala Pro Glu Ser Arg Gly Pro Ser Gly Ala Ser  
 20 25 30  
 Ala Leu Arg Cys Arg Gly Pro Thr Ala Arg Ser Leu Pro Ala Ala Ser  
 35 40 45  
 65 Met Leu Gly Ala Pro Asp Glu Ser Ser Val Arg Val Ala Val Arg Ile  
 50 55 60  
 Arg Pro Gln Leu Ala Lys Glu Lys Ile Glu Gly Cys His Ile Cys Thr

		65				70				75					80		
		Ser	Val	Thr	Pro	Gly	Glu	Pro	Gln	Val	Phe	Leu	Gly	Lys	Asp	Lys	Ala
						85					90					95	
5		Phe	Thr	Phe	Asp	Tyr	Val	Phe	Asp	Ile	Asp	Ser	Gln	Gln	Glu	Gln	Ile
					100					105						110	
		Tyr	Ile	Gln	Cys	Ile	Glu	Lys	Leu	Ile	Glu	Gly	Cys	Phe	Glu	Gly	Tyr
				115						120						125	
		Asn	Ala	Thr	Val	Phe	Ala	Tyr	Gly	Gln	Thr	Gly	Ala	Gly	Lys	Thr	Tyr
													140				
10		Thr	Met	Gly	Thr	Gly	Phe	Asp	Val	Asn	Ile	Val	Glu	Glu	Glu	Leu	Gly
							150						155				160
		Ile	Ile	Ser	Arg	Ala	Val	Lys	His	Leu	Phe	Lys	Ser	Ile	Glu	Glu	Lys
						165					170					175	
15		Lys	His	Ile	Ala	Ile	Lys	Asn	Gly	Leu	Pro	Ala	Pro	Asp	Phe	Lys	Val
					180						185					190	
		Asn	Ala	Gln	Phe	Leu	Glu	Leu	Tyr	Asn	Glu	Glu	Val	Leu	Asp	Leu	Phe
					195										205		
		Asp	Thr	Thr	Arg	Asp	Ile	Asp	Ala	Lys	Ser	Lys	Lys	Ser	Asn	Ile	Arg
									215						220		
20		Ile	His	Glu	Asp	Ser	Thr	Gly	Gly	Asn	Leu	Tyr	Cys	Gly	Arg	Phe	Gln
								230									240
		His	Val	Leu								235					

```
25      <210> 20
        <211> 223
        <212> PRT
        <213> Homo sapiens
```

30	<400> 20															
	Gly 1	Asn	Tyr	Arg 5	Pro	Ser	Lys	Lys	Arg 10	Phe	Arg	Glu	Gly	Lys	Glu 15	Lys
	Lys	Arg	Gly	Lys 20	Lys	Val	Trp	Leu	Val 25	Lys	Arg	Ile	Ile	Gln 30	Thr	Leu
35	Thr	Lys	Arg 35	Arg	Lys	Lys	Lys	Arg 40	Val	Phe	Arg	Lys	Glu 45	Lys	Pro	Asn
	Glu 50	Leu	Glu	Val	Glu	Glu	Ser 55	Gln	Glu	Val	Ser	Asp 60	His	Glu	Asp	Glu
40	Glu 65	Glu	Glu	Glu	Glu	Glu 70	Glu	Glu	Asp	Asp	Ile 75	Asp	Gly	Gly	Glu	Ser 80
	Ser	Asp	Glu	Ser	Asp 85	Ser	Glu	Ser	Asp	Glu 90	Lys	Ala	Tyr	Gln 95	Ala	Asp
	Leu	Ala	Asn	Ile 100	Thr	Cys	Glu	Ile	Ala 105	Ile	Lys	Gln	Lys	Leu 110	Ile	Asp
45	Glu	Leu	Glu	Asn 115	Ser	Gln	Lys	Arg	Leu 120	Gln	Thr	Leu	Lys 125	Lys	Gln	Tyr
	Glu	Glu	Lys	Met	Met	Leu	Gln 135	His	Lys	Ile	Arg	Asp 140	Thr	Gln	Leu	Glu
50	Arg 145	Asp	Gln	Val	Leu	Gln 150	Asn	Leu	Gly	Ser	Val 155	Glu	Ser	Tyr	Ser	Glu 160
	Glu	Lys	Ala	Lys 165	Lys	Val	Arg	Ser	Glu	Tyr 170	Glu	Lys	Lys	Leu 175	Gln	Ala
	Met	Asn	Lys	Glu 180	Leu	Gln	Arg	Leu	Gln 185	Ala	Ala	Gln	Lys	Glu 190	His	Ala
55	Arg	Leu	Leu 195	Lys	Asn	Gln	Ser	Gln 200	Tyr	Glu	Lys	Leu	Lys 205	Lys	Leu	Gln
	Gln	Asp 210	Val	Met	Glu	Met	Lys 215	Lys	Thr	Lys	Val	Arg	Leu 220	Met	Lys	

```
60      <210> 21
      <211> 206
      <212> PRT
      <213> Homo sapiens
```

```
65      <400> 21  
Met Asn Lys Pro Ile Thr Pro Ser Thr Tyr Val Arg Cys Leu Asn Val  
   1          5           10          15
```



	Gly	Leu	Ile	Arg	Lys	Leu	Ser	Asp	Phe	Ile	Asp	Pro	Gln	Glu	Gly	Trp	
			20						25				30				
	Lys	Lys	Leu	Ala	Val	Ala	Ile	Lys	Lys	Pro	Ser	Gly	Asp	Asp	Arg	Tyr	
			35					40					45				
5	Asn	Gln	Phe	His	Ile	Arg	Arg	Phe	Glu	Ala	Leu	Leu	Gln	Thr	Gly	Lys	
		50				55						60					
	Ser	Pro	Thr	Ser	Glu	Leu	Leu	Phe	Asp	Trp	Gly	Thr	Thr	Asn	Cys	Thr	
	65				70					75					80		
10	Val	Gly	Asp	Leu	Val	Asp	Leu	Leu	Ile	Gln	Asn	Glu	Phe	Phe	Ala	Pro	
				85					90						95		
	Ala	Ser	Leu	Leu	Leu	Pro	Asp	Ala	Val	Pro	Lys	Thr	Ala	Asn	Thr	Leu	
			100					105						110			
	Pro	Ser	Lys	Glu	Ala	Ile	Thr	Val	Gln	Gln	Lys	Gln	Met	Pro	Phe	Cys	
		115					120						125				
15	Asp	Lys	Asp	Arg	Thr	Leu	Met	Thr	Pro	Val	Gln	Asn	Leu	Glu	Gln	Ser	
	130					135						140					
	Tyr	Met	Pro	Pro	Asp	Ser	Ser	Ser	Pro	Glu	Asn	Lys	Ser	Leu	Glu	Val	
	145				150						155				160		
20	Ser	Asp	Thr	Arg	Phe	His	Ser	Phe	Ser	Phe	Tyr	Glu	Leu	Lys	Asn	Val	
				165						170					175		
	Thr	Asn	Asn	Phe	Asp	Glu	Arg	Pro	Ile	Ser	Val	Gly	Gly	Asn	Lys	Met	
			180						185					190			
	Gly	Glu	Gly	Gly	Phe	Gly	Val	Val	Tyr	Lys	Gly	Leu	Arg	Lys			
		195					200						205				
25																	
	<210>	22															
	<211>	1260															
	<212>	DNA															
	<213>	Homo sapiens															
30																	
	<400>	22															
	cttctccgca	cgaactgttac	agagggtcttc	agagccttct	ctctcctgtg	caaaaatggca		60									
	actcttaagg	aaaaactcat	tgcaccagtt	gcggaagaag	aggcaacagt	tccaacaat		120									
35	aagatcactg	tagtggtgtg	tggacaagtt	ggtatggcgt	gtgctatcag	cattctggga		180									
	aagtctctgg	ctgatgaact	tgctcttgtg	gatgttttgg	aagataagct	taaaggagaa		240									
	atgatggatc	tcgacgatgg	gagcttattt	cttcagacac	ctaaaattgt	ggcgataaaa		300									
	gattattctg	tgaccgccaa	ttctaagatt	gtagtggtaa	ctgcaggagt	ccgtcagcaa		360									
	gaaggggaga	gtcggctcaa	tctggtgcag	agaaatgta	atgtcttcaa	attcattatt		420									
40	cctcagatcg	tcaagtacag	tctgtattgc	atcataattg	tggtttccaa	cccagtggac		480									
	attcttagct	atgttacctg	gaaactaaat	ggattaccac	aacaccgcgt	gattggaagt		540									
	ggatgtaate	tggattctgc	tagatttctg														

	atcaacaccg	aggtgctgag	cttcggacct	gtgtcccga	gctgtcttca	cctggccttc	540
	caggactatg	gctgctgcat	gtccctcatc	gccgtgctg	tcttctaccg	caagtgtccc	600
	cgcacatccc	agaatggcgc	catcttccag	gaaaccctgt	cggtggctga	gagcacatcg	660
5	ctgtgtgctg	cccggggcag	ctgcatcgcc	aatgctggaag	aggtggatgt	acccatcaag	720
	ctctactgta	acggggacgg	cgagtggctg	gtgcccacg	ggcgtgcat	gtgcaaagca	780
	ggcttcgagg	ccgttgagaa	tggcaccgtc	tgcgaggtt	gtccatctgg	gactttcaag	840
	gccaaccaag	gggatgaggg	ctgtaccac	tgtcccatca	acagccggac	cactttctgaa	900
	ggggccacca	actgtgtctg	ccgcaatggc	tactacagag	cagacctgga	ccccctggac	960
	atggccctgca	caaccatccc	ctccgcgcc	caggctgtga	tttccagtgt	caatgagacc	1020
10	tcctcatg	tggagtggac	ccctcccgc	gactccggag	gccgagagga	cctcgtctac	1080
	aacatcatct	gcaagagctg	tggctcgggc	cggtgtgcct	gcacccgtg	cggtgacaat	1140
	gtacagtacg	caccacgcca	gctaggcctg	accgagccac	gcatttacat	cagtgcacctg	1200
	ctggccacaca	cccagtagac	cttcgagatc	caggctgtga	acggcggtac	tgaccagagc	1260
	cccttctcgc	ctcagttcgc	ctctgtgaac	atcaccacca	accaggcagc	tccatcggca	1320
15	gtgtccatca	tgcacaggt	gagccgcacc	gtggacagca	ttaccctgtc	gtgttccag	1380
	ccagaccagc	ccaatggcgt	gatcctggac	tatgagctgc	agtactatga	gaagcaggag	1440
	ctcagttagt	acaacgccac	agccataaaa	agccccacca	acacggtcac	cgtgcagggc	1500
	ctcaaagccg	gcgcatctca	tgtcttccag	gtgcgggcac	gcacccgtggc	aggctacggg	1560
20	cgctacagcg	gcaagatgta	cttccagacc	atgacagaag	ccgattacca	gacaagcatc	1620
	caggagaagt	tgccactcat	catcggtccc	tcggccgctg	gcctgtgtctt	cctcattgct	1680
	gtggttgtca	tcgccatcgt	gtgtaacaga	cggtgtgttg	agcgtgctga	ctcggagtac	1740
	acggacaagc	tgcaacacta	caccagtggc	cacatgaccc	caggcatgaa	gatctacatc	1800
	gaccccttca	cctacagagga	ccccaacgag	cgagtgcggg	agtttgccaa	ggaaattgac	1860
25	atctcctgtg	tcaaaattga	gcaggtgatc	ggagcagggg	agtttgccga	ggtctgcagt	1920
	ggccacctga	agctgccagg	caagagagag	atctttgtgg	ccatcaagac	gctcaagtcg	1980
	ggctacacgg	agaagcagcg	ccgggacttc	ctgagcgaag	cctccatcat	gggccagttc	2040
	gacccatccc	acgtcatcca	cctggagggt	gtcgtgacca	agagcacacc	tgtgatgatc	2100
	atcaccgagt	tcatggagaa	tggctccctg	gactcctttc	tccggcaaaa	cgtatggcag	2160
30	ttcacagtca	tccagctggt	gggcatgctt	cggtgacatc	cagctggcat	gaagtacctg	2220
	gcagacatga	actatgttca	ccgtgacctg	gctgcccga	acatcctcgt	caacagcaac	2280
	ctggtctgca	aggtgtcggg	ctttgggctc	tcacgctttc	tagaggacga	tacctagac	2340
	cccaccta	ccagtgcctt	gggcggaaaag	atccccatcc	gctggacagc	ccggaagcc	2400
	atccagtacc	ggaagtac	ctcggccagt	gatgtgtgga	gctacggcat	tgtcatgtgg	2460
	gaggtgatgt	cctatgggga	gcggccctac	tgggacatga	ccaaccagga	tgtaatcaat	2520
35	gccattgagc	aggactatcg	gctgccaccg	cccatggact	gcccagctgc	cctgcaccaa	2580
	ctcatgctgg	actgttggca	gaagaccggc	aaccaccggc	ccaagttcgg	ccaaattgtc	2640
	aacacgctag	acaagatgat	ccgcaatccc	aacagcctca	aagccatggc	gcccctctcc	2700
	tctggcatca	acctgccgct	gctggaccgc	acgatccccg	actacaccag	ctttaacacg	2760
	tgggacaggt	ggctgaaggc	catcaagatg	gggcagtaca	aggagagctt	cgccaatgcc	2820
40	ggcttcaact	cccttgacgt	cgtgtctcag	atgatgatgg	aggacattct	ccgggttggg	2880
	gtcactttgg	ctggccaaca	gaaaaaaatc	ctgaaacagta	tccaggtgat	gcgggcgcag	2940
	atgaaccaga	ttcagctctg	ggaggtttga	cattcaactg	cctcggctca	cctcttctc	3000
	caagccccgc	cccctctgcc	ccacgtgccg	gccctcctgg	tgtctctatcc	actgcagggc	3060
	cagccactcg	ccaggaggcc	acgggcacgg	gaagaaccaa	gcggtgccag	ccacgagacg	3120
45	tcaccaagaa	aacatgcaac	tcaaacgacg	g			3151

&lt;210&gt; 24

&lt;211&gt; 1234

&lt;212&gt; DNA

50 &lt;213&gt; Homo sapiens

&lt;400&gt; 24

	tagttcaaga	caacagagac	aaagctaaga	tgaggaaagt	ctgtacagtt	taggaaatag	60
	aggctttcaa	agataattcg	cagtgtatgtg	aaactggcct	cccaagccct	gataacaaca	120
55	tggccaacgc	cctggccagc	gccacttgcg	agcgtgcaa	gggcggcttt	gcgcccgtg	180
	agaagatcgt	gaacagtaat	ggggagctgt	accatgagca	gtgtttcgtg	tgcgctcagt	240
	gcttccagca	gttcccagaa	ggactcttct	atgagtttga	aggaagaaag	tactgtgaac	300
	atgactttca	gatgctcttt	gccccttgct	gtcatcagtg	tggatgaattc	atcattggcc	360
	gagttatcaa	agccatgaat	aacagctggc	atccggagtg	cttccgctgt	gacctctgcc	420
60	aggaagttct	ggcagatata	gggtttgtca	agaatgctgg	gagacacctg	tgtcgccctt	480
	gtcataatcg	tgagaaagcc	agaggccttg	ggaaatacat	ctgcagaaa	tgccatgcta	540
	tcacgatga	gcagcctctg	atattcaaga	acgaccctca	ccatccagac	catttcaact	600
	gcgccaactg	cggaagagag	ctgactgccg	atgcacggga	gctgaaaggg	gagctatact	660
	gcctcccatg	ccatgataaa	atgggggtcc	ccatctgtgg	tgttgccga	cgccccatcg	720
65	aagggcgctg	ggtgaacgct	atgggcaagc	agtggcatgt	ggagcatttt	gtttgtgcca	780
	agtgtgagaa	accctttctt	ggacatcgcc	attatgagag	gaaaggcctg	gcatattgtg	840
	aaactcacta	taaccagcta	tttgggtgatg	tttgcctcca	ctgcaatcgt	gttatagaag	900

	gtgatgtggt	ctctgctctt	aataaggcct	ggtgctgaa	ctgctttgcc	tgtttacct	960
	gcaacactaa	attaacactc	aagaataagt	ttgtggagtt	tgacatgaag	ccagtcgtga	1020
	agaagtgc	tgagatttcc	attggagctg	aagaaaagac	ttaagaaact	agctgagacc	1080
	ttaggaagga	aataagttcc	tttatttttt	cttttctatg	caagataaga	gattaccaac	1140
5	attacttgtc	ttgatctacc	catattttaa	gctatatctc	aaagcagttg	agagaagagg	1200
	acctatatga	atgggtttat	gtcatttttt	taaa			1234
	<210> 25						
	<211> 4534						
10	<212> DNA						
	<213> Homo sapiens						
	<400> 25						
15	gtcacgagcg	tcgaagagac	aaagccgcgt	cagggggccc	ggccggggcg	ggggagcccc	60
	gggcttgttg	gtgccccagc	ccgcgcggag	ggcccttcgg	acccgcgcgc	cgccgctgcc	120
	gccgccgcgc	cctcgcaaca	ggtccgggag	gctcgcctct	ccgctccctc	ccccgcctc	180
	cgcgaccctc	cggggcacct	cagctcggcc	ggggccgcag	tctggccacc	cgttccatg	240
	cggttcgggt	ccaagatgat	gccgatgttt	cttaccgtgt	atctcagtaa	caatgagcag	300
20	cacttcacag	aagtccag	tactccagaa	acaatatgca	gagacgtggt	ggatctgtgc	360
	aaagaacccg	gcgagagtga	ttgccatttg	gctgaagtgt	gggtgtggctc	tgtagagata	420
	gagtttcatc	atgttggcca	ggatgggtctc	gatctcctga	ccttgtgac	cgctgcctc	480
	ggcctcccaa	agtctggat	tacaggtgtg	agccaccacg	atcagcctct	agtgtttaa	540
	aaagaacgtc	cagtctcgga	taattgagca	atgtttgatg	ttcttcaacg	atttggaggt	600
25	cagaggaacg	aagtctcgctt	cttccttcgt	catgaacgcc	cccctggcag	ggacattgtg	660
	agtggaccaa	gatctcagga	tccaagttta	aaaagaaatg	gtgtaaaagt	tcctggtgaa	720
	tatcgaagaa	aggagaacgg	tgtaaatagt	cctaggatgg	atctgactct	tgctgaactt	780
	caggaaatgg	catctcgcca	gcagcaacag	attgaagccc	agcaacaatt	gctggcaact	840
	aaggaacagc	gcttaaagtt	tttgaaacaa	caagatcagc	gacaacagca	acaagttgct	900
30	gagcaggaga	aacttaaaa	gctaaaagaa	atagctgaga	atcaggaagc	taagctaaaa	960
	aaagtgagag	cacttaaaag	ccacgtggaa	cagaagagac	taagcaatgg	gaaacttgtg	1020
	gaggaaattg	aacagatgaa	taatttgttc	cagcaaaaac	agagggagct	cgtcctggct	1080
	gtgtcaaaa	tagaagaact	gaccaggcag	ctagagatgc	tcaagaacgg	caggatcgac	1140
	agccaccatg	acaatcagtc	tgacgtggct	gagcttgatc	gcctctataa	ggagctgcag	1200
	ctaagaaaca	aattgaatca	agagcagaat	gccaaagctac	aacaacagag	ggagtgtttg	1260
35	aataagcgta	attcagaagt	ggcagtcacg	gataagcgtg	ttaatgagct	gagggaccgg	1320
	ctgtggaaga	agaaggcagc	tctacagcaa	aaagaaaatc	taccagtttc	atctgatgga	1380
	aatcttcccc	agcaagccgc	gtcagcccca	agccgtgtgg	ctgcagtagg	tccttatatc	1440
	cagtcaccta	ctatgcctcg	gatgccctca	aggcctgaat	tgctggtgaa	gccagccctg	1500
	ccggatgggt	ccttgggtcat	tcaggcttca	gaggggccga	tgaaaataca	gacactgccc	1560
40	aacatgagat	ctggggctgc	ttcacaaact	aaaggctcta	aaatccatcc	agttggccct	1620
	gattggagtc	cttcaaatgc	agatcttttc	ccaagccaag	gctctgcttc	tgtacctcaa	1680
	agcactggga	atgctctgga	tcaagttgat	gatggagagg	ttccgctgag	ggagaaagag	1740
	aagaaagtgc	gtccgttctc	aatgtttgat	gcagttagacc	agtccaatgc	cccaccttcc	1800
	tttggtaact	tgaggaagaa	ccagagcagt	gaagatatct	tgccggatgc	tcaggttgca	1860
45	aataaaaaatg	tggttaaaatg	accacctcct	gttctacaaa	aaccaaaaca	gattaatttg	1920
	ccttattttg	gacaaactaa	tcagccacct	tcagacatta	agccagacgg	aagttctcag	1980
	cagttgtcaa	cagttgttcc	gtccatggga	actaaaccaa	aaccagcagg	gcagcagccg	2040
	agagtgtctc	tatctcccag	cataccttcc	gttggccaa	accagaccct	ttctccaggt	2100
	tctaagcaag	aaagttccac	tgctgtgtgc	gtccggccct	ttactcccca	gccttccaaa	2160
50	gacaccttac	ttccaccctt	cagaaaaccc	cagaccgtgg	cagcaagttc	aatatatattc	2220
	atgtatacgc	aacagcaggc	gccaggaaaa	aacttccagc	aggctgtgca	gagcgcgttg	2280
	accaagactc	ataccagagg	gccacacttt	tcaagtgtat	atggtaagcc	tgtaattgct	2340
	gctgccaga	atcaacagca	gcaccagag	aacatttatt	ccaatagcca	gggcaagcct	2400
	ggcagtcag	aacctgaaac	agagcctggt	tcttcagttc	aggagaacca	tgaaaacgaa	2460
55	agaattcctc	ggccactcag	cccaactaaa	ttactgcctt	tcttatctaa	tccttaccga	2520
	aaccagagtg	atgctgacct	agaagcctta	cgaagaaac	tgtctaacgc	accaaggcct	2580
	ctaaagaaac	gtagttctat	tacagagcca	gagggctcta	atgggcca	tattcagaag	2640
	cttttatatc	agaggaccac	catagcggcc	atggagacca	tctctgtccc	atcataccca	2700
	tccaagtccag	cttctgtgac	tgccagctca	gaagcccgag	tagaaatcca	gaatccat	2760
60	ttacatgtgg	agcccga	ggaggtgggc	tctctgtgtc	ctgaatcatt	gtccccagag	2820
	gatgtgggga	atgccagtac	agagaacagt	gacatgccag	ctccttctcc	aggccttgat	2880
	tatgagcctg	aggaggtccc	agacaacagc	ccaaatctcc	agaataaccc	agaagaacca	2940
	aatccagagg	ctccacatgt	gcttgatgtg	tacttgagg	agtaccctcc	ataccaccc	3000
	ccaccatacc	catctgggga	gcctgaaggg	cccggagaag	actcggtgag	catgcgccc	3060
65	cctgaaatca	ccgggcaggt	ctctctgcct	cctggtaaaa	ggacaaactt	gcgtaaaact	3120
	ggctcagagc	gtatcgctca	tggaatgagg	gtgaaattca	accccttgc	tttactgcta	3180
	gattcgtctt	tggaaggaga	atttgacctt	gtacagagaa	ttatttatga	ggttgatgac	3240

	ccaagcctgc	ccaatgatga	aggcatcacg	gctcttcaca	atgctgtgtg	tgcaggccac	3300
	acagaaatcg	ttaagttcct	ggtacagttt	ggtgtaaattg	taaatgctgc	tgatagtgat	3360
	ggatggactc	cattacattg	tgctgcctca	tgtaacaacg	tccaagtgtg	taagtttttg	3420
5	gtggagtcag	gagccgctgt	gtttgccatg	acctacagt	acatgcagac	tgctgcagat	3480
	aagtgcgagg	aaatggagga	aggctacact	cagtgtctcc	aatttcttta	tggagttcag	3540
	gagaagatgg	gcataatgaa	taaaggagtc	atztatgcgc	tttgggatta	tgaacctcag	3600
	aatgatgatg	agctgcccac	gaaagaagga	gactgcatga	caatcatcca	cagggaagac	3660
	gaagatgaaa	tcgaatggtg	gtgggcgcg	cttaatgata	aggagggata	tgttccacgt	3720
10	aacttgctgg	gactgtaccc	aagaattaaa	ccaagacaaa	ggagcttggc	ctgaaacttc	3780
	cacacagaat	tttagtcaat	gaagaattaa	tctctgttaa	gaagaagtaa	tacgattatt	3840
	tttggcaaaa	atttcacaag	acttattttta	atgacaatgt	agcttgaaag	cgatgaagaa	3900
	tgctctaga	agagaatgaa	ggattgaaga	attcaccatt	agaggacatt	tagcgtgatg	3960
	aaataaagca	tctacgtcag	caggccatac	tgtgttgggg	caaagggtgc	ccgtgtagca	4020
	ctcagataag	tatacagcga	caatcctggt	ttctacaaga	atcctgtcta	gtaaatagga	4080
15	tcattttattg	ggcagttggg	aaatcagctc	tctgtcctgt	tgagtgtttt	cagcagctgc	4140
	tcctaaacca	gtcctcctgc	cagaaaggac	cagtgcctgc	acatcgctgt	ctctgattgt	4200
	ccccggcacc	agcaggcctt	ggggctcact	gaaggctcga	aggcactgca	caccttgtat	4260
	attgtcagtg	aagaacgtta	gttggttgtc	agtgaacaat	aactttatta	tatgagtttt	4320
	tgtagcatct	taagaattat	acatatgttt	gaaatattga	aactaagcta	cagtaccagt	4380
20	aattagatgt	agaatcttgt	ttgtaggctg	aattttaatc	tgtatttatt	gtcttttgta	4440
	tctcagaaat	tagaaacttg	ctacagactt	acccgtaata	tttgtcaaga	tcatagctga	4500
	ctttaaaaac	agttgtaata	aactttttga	tgct			4534
25	<210> 26						
	<211> 4660						
	<212> DNA						
	<213> Homo sapiens						
	<400> 26						
30	ggggccttaga	aaataacagg	ttgtttatat	aattggcctt	aaatgaggtg	agagtgaaga	60
	gactagagcc	atctctggaa	aacatcatta	tccatttccc	cggaagcta	ccctctggaa	120
	ctcaagattt	gaccatatct	gttttgagga	ttcattatga	acaaagaagt	ctcccaggtg	180
	tgaagttttt	caacatgagt	ggcctcgagg	acagttcatc	cgaccctgct	aaccagact	240
	caataagag	gaaaggatcg	ccatgtgaca	cactggcatc	aagcacggaa	aagaggcgca	300
35	gggagcaaga	aaataaata	ttagaagaac	tagctgagtt	actgtctgcc	aacattagtg	360
	acattgacag	cttgagtgtg	aaaccagaca	aatgcaagat	tttgaagaaa	acagtcgac	420
	agatacagct	aatgaagaga	atggaacaag	agaaatcaac	aactgatgac	gatgtacaga	480
	aatcagacat	ctcatcaagt	agtcaaggag	tgatagaaaa	ggaatccttg	ggaccccttc	540
	ttttggaggc	tttggatgga	ttttctgttg	ttgtgaactg	tgaagggaga	attgtatttg	600
40	tgtagagaga	tgtaaccagc	tacttaggtt	acaatcagga	ggaattaatg	aataaccagc	660
	tctacagcat	actgcacgtg	gggatcatg	cagaatttgt	gaagaatctg	ctacccaaat	720
	cactagtaaa	tggagtccct	tggcctcaag	aggcaacacg	acgaaatagc	cataccttta	780
	actgcaggat	gctaattcac	cctccagatg	agccagggac	cgagaaccaa	gaagcttgcc	840
	agcgttatga	agtaatgcag	tgtttcaact	tgccacagcc	aaaatcaatt	caagaggatg	900
45	gagaagattt	ccagtcattg	ctgatttgta	ttgcacggcg	attacctcgg	cctccagcta	960
	ttacgggtgt	agaatccttt	atgaccaagc	aagatactac	aggtaaaatc	atctctattg	1020
	atactagttc	cctgagagct	gctggcagaa	ctggttggga	agatttagtg	aggaaagtga	1080
	tttatgcttt	tttccaacct	cagggcagag	aacctctta	tgccagacag	ctgttccaag	1140
	aagtgtgac	tcgtggcact	gcctccagcc	cctcctatag	attcatattg	aatgatggga	1200
50	caatgcttag	cgcccacacc	aagtgtaaac	tttgctaccc	tcaaagtcca	gacatgcaac	1260
	ctttcatcat	gggaattcat	atcatcgaca	gggagcacag	tgggctttct	cctcaagatg	1320
	acactaattc	tggaatgtca	attccccgag	taaatccctc	ggtcaatcct	agtatctctc	1380
	cagctcatgg	tgtggctcgt	tcatccacat	tgccaccatc	caacagcaac	atggtatcca	1440
	ccagaataaa	ccgccagcag	agctcagacc	ttcatagcag	cagtcatagt	aattctagca	1500
55	acagccaagg	aagtttcgga	tgctcaccgg	gaagtcagat	tgtagccaat	gttgcccttaa	1560
	accaaggaca	ggccagttca	cagagcagta	atccctcttt	aaacctcaat	aattctccta	1620
	tggaaaggta	aggaatatcc	ctagcacagt	tcatgtctcc	aaggagacag	gttacttctg	1680
	gattggcaac	aaggcccgag	atgccaacaa	attcctttcc	tcctaataat	tcgacattaa	1740
	gctctcccg	tggcatgaca	agtagtgctt	gtaataataa	taaccgatct	tattcaaaca	1800
60	tcccagtaac	atctttacag	ggtatgaatg	aaggacccaa	taactccgtt	ggcttctctg	1860
	ccagttcttc	agtcctcagg	cagatgagct	cacagaattc	acctagcaga	ttaaataatac	1920
	aaccagcaaa	agctgagtc	aaagataaca	aagagattgc	ctcaatttta	aatgaaatga	1980
	ttcaatctga	caacagctct	agtgatggca	aacctctgga	ttcagggtct	ctgcataaca	2040
	atgacagact	ttcagatgga	gacagtaaat	actctcaaac	cagtcacaaa	ctagtgcagc	2100
65	ttttgacaac	aactgccgaa	cagcagttac	ggcatgctga	tatagacaca	agctgcgaag	2160
	atgtcctgtc	ttgcacagg	acttccaact	ctgcctctgc	taactcttca	ggagggtctt	2220
	gtccctcttc	tcatagctca	ttgacagaac	ggcataaaat	tctacaccgg	ctcttacagg	2280

	agggtagccc	ctcagatatc	accactttgt	ctgtcgagcc	tgataaaaag	gacagtgcac	2340
	ctactttctgt	gtcagtgact	ggacaggtac	aaggaaactc	cagtataaaa	ctagaactgg	2400
	atgcttcaaa	gaaaaaagaa	tcaaaagacc	atcagctcct	acgctatctt	ttagataaaag	2460
	atgagaaaaga	tttaagatca	actccaaacc	tgagcctgga	tgatgtaaag	gtgaaagtgg	2520
5	aaaagaaaaga	acagatggat	ccatgttaata	caaacccaac	cccaatgacc	aaaccactc	2580
	ctgaggaaat	aaaactggag	gcccagagcc	agtttacagc	tgaccttgac	cagtttgatc	2640
	agttactgcc	cacgctggag	aaggcagcac	agttgccagg	cttatgtgag	acagacagga	2700
	tggtatgtgc	ggtcaccagt	gtaaccatca	aatcggagat	cctgccagct	tcacttcagt	2760
	ccgccactgc	cagaccctc	tccaggctaa	atagattacc	tgagctggaa	ttggaagcaa	2820
10	ttgataacca	atgttgacaa	ccaggaacag	gcgctcagat	tccatggaca	aataatcacg	2880
	tgacagctat	aaatcagagt	aaatcagaag	accagtgtat	tagctcacia	ttagatgagc	2940
	ttctctgtcc	acccacaaca	gtagaaggga	gaaatgatga	gaaggctctt	cttgaacagc	3000
	tggtatcctt	ccttagtggc	aaagatgaaa	ctgagctagc	tgaactagac	agagctctgg	3060
	gaattgacaa	acttgttcag	gggggtggat	tagatgtatt	atcagagaga	tttccaccac	3120
15	aacaagcaac	gccacctttg	atcatggaag	aaagacccaa	cctttattcc	cagccttact	3180
	cttctccttc	tcctactgcc	aatctcccta	gccctttcca	aggcatggc	aggcaaaaac	3240
	cttcaactgg	gacgatgcct	gttcaagtaa	cacctccccg	aggtgctttt	tcacctggca	3300
	tgggcatgca	gcccaggcaa	actctaaaca	gacctccggc	tgacaccta	cagcttcgac	3360
	ttcaactaca	gcagcgatta	cagggacaac	agcagttgat	acaccaaaa	cgcaagcta	3420
20	tcttaaacca	gtttgcagca	actgctcctg	ttggcatcaa	tatgagatca	ggcatgcaac	3480
	agcaaattac	acctcagcca	ccctgaatg	ctcaaagtgt	ggcacaacgt	cagcggaac	3540
	tgtacagtca	acagcaccca	cagaggcagc	taatacagca	gcaaagagcc	atgcttatga	3600
	ggcagcaaa	ctttgggaac	aacctccctc	cctcatctgg	actaccagtt	caaagtggga	3660
	acccccgtct	tcctcaggg	gctccacagc	aattccctca	tccaccaaac	tatggtacaa	3720
25	atccaggaac	cccacctgct	tctaccagcc	cgttttcaca	actagcagca	aatcctgaag	3780
	catccttggc	caaccgcaac	agcatggtga	gcagaggcat	gacaggaaac	ataggaggac	3840
	agtttggcac	tggaaatcaat	cctcagatgc	agcagaatgt	cttccagtat	ccaggagcag	3900
	gaatggttcc	ccaaggtgag	gccaaacttg	ctccatctct	aagccctggg	agctccatgg	3960
	tgccgatgcc	aatccctcct	cctcagagtt	ctctgctcca	gcaaactcca	cctgcctccg	4020
30	ggtatcagtc	accagacatg	aaggcctggc	agcaaggagc	gataggaaac	aacaatgtgt	4080
	tcagtcgaag	tgtccagaac	cagcccacgc	ctgcacagcc	aggagtatac	aacaacatga	4140
	gcatcacctg	ttccatggca	ggtggaaata	cgaatgttca	gaacatgaac	ccaatgatgg	4200
	cccagatgca	gatgagctct	ttgcagatgc	caggaatgaa	caactgtgtg	cctgagcaga	4260
	taaatgatcc	cgcactgaga	cacacaggcc	tctactgcaa	ccagctctca	tccactgacc	4320
35	ttctcaaaa	agaagcagat	ggaaccagc	tgcaacaggt	tcagggtgtt	gctgacgtcc	4380
	agtgtacagt	gaatctggta	ggcggggacc	cttactgtaa	ccagctgggt	ccactgggaa	4440
	ctcaaaagcc	cacgtcagga	ccacagaccc	cccaggccca	gcagaagagc	ctccttcagc	4500
	agctactgac	tgaataacca	cttttaaaag	aatgtgaaat	ttaaataata	gacatacaga	4560
	gatatacaaa	tatattatat	atttttctga	gatttttgat	atctcaatct	gcagccattc	4620
40	ttcaggtcgt	agcatttggg	gcaaaaaaaa	aaaaaaaaaa			4660

&lt;210&gt; 27

&lt;211&gt; 6773

&lt;212&gt; DNA

45 &lt;213&gt; Homo sapiens

&lt;400&gt; 27

	gcggtctggt	gcgggccggc	ggcgggctgg	cggagatgga	ggatcttgtt	caagatgggg	60
	tggtcttacc	agctaccctt	gggaccggga	aatctaagaa	ttggagaaag	aaattgaaga	120
50	actcagatca	aaacctgtta	ctgaaggaa	tggtgatatt	attaaggcat	taactgaacg	180
	tctggatgct	cttcttctgg	aaaaagcaga	gactgagcaa	cagtgtcttt	ctctgaaaaa	240
	ggaaaatata	aaaatgaagc	aagaggttga	ggattctgta	acaaagatgg	gagatgcaca	300
	taaggagttg	gaacaatcac	atataaacta	tgtgaaagaa	attgaaaatt	tgaaaaatga	360
	gttgatggca	gtacgtttcca	aatacagtga	agacaaaagc	aacttacaaa	agcagctgga	420
55	agaagcaatg	aatacgcaat	tagaactttc	agaacaactt	aaatttcaga	acaactctga	480
	agataatgtt	aaaaaactac	aagaagagat	tgagaaaatt	aggccaggct	ttgaggagca	540
	aattttatat	ctgcaaaaagc	aattagacgc	taccactgat	gaaaagaagg	aaacagttac	600
	tcaactccaa	aatatcattg	aggctaattc	tcagcattac	caaaaaata	ttaatagttt	660
	gcaggaagag	cttttacagt	tgaagctat	acaccaagaa	gaggtgaaag	agttgatgtg	720
60	ccagattgaa	gcatcagcta	aggaacatga	agcagagata	aataagttga	acgagctaaa	780
	agagaactta	gtaaaacaat	gtgaggcaag	tgaagaagaa	atccagaaga	aatatgaatg	840
	tgagttagaa	aatttaagga	aagccacctc	aatgcaaac	caagacaatc	agatatgttc	900
	tattctcttg	caagaaaaata	catttttaga	acaagtagta	aatgaaaaag	tcaaacactt	960
	agaagatacc	ttaaaagaac	ttgaatctca	acacagtatc	ttaaaagatg	aggtaactta	1020
65	tatgaataat	cttaagttaa	aacttgaaat	ggatgctcaa	catataaagg	atgagttttt	1080
	tcatgaacgg	gaagacttag	agtttaaaat	taatgaatta	ttactagcta	aagaagaaca	1140
	gggctgtgta	attgaaaaat	taaaatctga	gctagcaggt	ttaaataaac	agttttgcta	1200

	tactgtagaa	cagcataaca	gagaagtaga	gagtccttaag	gaacaacatc	aaaaagaaat	1260
	atcagaacta	aatgagacat	ttttgtcaga	ttcagaaaaa	gaaaaattaa	cattaatgtt	1320
	tgaataacag	ggtcttaagg	aacagtgtaga	aaacctacag	caagaaaagc	aagaagcaat	1380
	tttaaattat	gagagtttac	gagagattat	ggaaatttta	caaacagaa	tgggggaatc	1440
5	tgctggaaaa	ataagtcaag	agttcgaatc	aattgaagcaa	cagcaagcat	ctgatgttca	1500
	tgaactgcag	cagaagctca	gaactgcttt	tactgaaaaa	gatgcccttc	tcgaaactgt	1560
	gaatcgccctc	cagggagaaa	atgaaaagtt	actatctcaa	caagaattgg	taccagaact	1620
	tgaaaatacc	ataaagaacc	ttcaagaaaa	gaatggagta	tacttactta	gtctcagtca	1680
	aagagatacc	atgtttaaag	aattagaagg	aaagataaat	tctcttactg	agggaaaaaga	1740
10	tgattttata	aataaactga	aaaattccca	tgaagaaatg	gataatttcc	ataagaaatg	1800
	tgaagggaag	gaaagattga	ttcttgaact	tgggaagaaa	gtagagcaaa	caatccagta	1860
	caacagtgaa	ctagaacaaa	aggtaaatga	attaacagga	ggactagagg	agactttaaa	1920
	agaaaaggat	caaaatgacc	aaaaactaga	aaaacttatg	gttcaaatga	aagttctctc	1980
	tgaagacaaa	gaagtattgt	cagctgaagt	gaagtctctt	tatgaggaaa	acaataaact	2040
15	cagttcagaa	aaaaaacagt	tgagtaggga	tttgagggtt	tttttgtctc	aaaaagaa	2100
	tggtatccctt	aaagaacata	ttactcaatt	agaaaagaaa	cttcagttaa	tgggtgaaga	2160
	gcaagataat	ttaaataaac	tgcttgaaaa	tgagcaagtt	cagaagttat	ttgttaaaac	2220
	tcagttgtat	ggtttttctt	aagaaaatggg	atcagaagtt	tcagaagaca	gtgaagagaa	2280
	agatgttggt	aatgtcctac	aggcagtcgg	tgaatccttg	gcaaaaaata	atgaggaaaa	2340
20	atgcaacctg	gcttttcagc	gtgatgaaaa	agtattagag	ttagaaaaag	agattaagtg	2400
	ccttcaagaa	gagagtgtag	ttcagtgtaga	agaacttaag	tctttattga	gagactatga	2460
	gcaagagaaa	gttctcttaa	ggaaagagtt	agaagaaata	cagtcagaaa	aagaggccct	2520
	gcagtcgtat	ctcttagaaa	tgaagaatgc	taatgaaaaa	acaaggcttg	aaaatcagaa	2580
	tcttttaatt	caagttgaag	aagtatctca	aacatgtagc	aaaagtgaag	tccataatga	2640
25	aaaagaaaaa	tgttttataa	aggaacatga	aaacctaaag	ccactactag	aacaaaaaga	2700
	attacgagat	aggagagcag	agttgatact	attaaaggat	tccttagcaa	aatcaccttc	2760
	tgtaaaaaat	gattcctctgt	cttcagttaa	agagttggaa	gaaaaaatag	aaaatctgga	2820
	aaaagaatgc	aaagaaaagg	aggagaaaa	aaataagata	aaattagttg	ccgtaaaagg	2880
	aaagaaagaa	ctagattcca	gcagaaaaga	gacccagact	gtgaagggaag	aacttgaatc	2940
30	tcttcgatca	gaaaaggacc	agttatctgc	ttccatgaga	gatctcatc	aaggagcaga	3000
	aagctataag	aatctttttat	tagaatatga	aaagcagtc	gagcaactgg	atgtggaaaa	3060
	agaacgtgct	aataattttg	agcatcgtat	tgaagacctt	acaagacaat	taagaaattc	3120
	gactttgcag	tgtgaacaa	taaattctga	taatgaagat	ctcctggctc	gtattgagac	3180
	attacagctc	aatgccaat	tattagaagt	acagatttta	gaagtccaga	gagccaaagc	3240
35	aatgtagagc	aaagaattag	aagctgaaaa	acttcagaaa	gaacagaaga	taaaaggaac	3300
	tgccactact	gtaaaatgaac	ttgaagaact	ctcaggtaca	cttcaaaagg	aaaagaaca	3360
	gcttcagaaa	accatgcaag	aattagagct	ggttaaaaag	gatgcccaac	aaaccacatt	3420
	gatgaatatg	gaaatagctg	attatgaacg	tttgatgaaa	gaactaaatc	aaaagttaac	3480
	taataaaaaac	aacaagatag	aagatttgga	gcaagaaata	aaaattcaaa	aacagaaaca	3540
40	agaaacccta	caagaagaaa	taacttcatt	acagtcctca	gtacaacaat	atgaagaaaa	3600
	aaacacccaa	atcaagcaat	tgcttgtaga	aaccaaaaag	gaactggcag	attcaagaca	3660
	agcagaaact	gatcacttaa	tacttcaagc	atcttttaaa	ggtagagctg	aggcaagcca	3720
	gcagcaagta	gaagtctata	aaatacagct	ggctgaaata	acatcagaga	agcacaacaa	3780
	ccacgagcac	ctgaaaacct	ctgcggaaca	gcaccagcgt	acgctaagtg	cataccagca	3840
45	gagagtgaca	gcactacagg	aagagtgccg	tgctgcgaag	gcagaacaag	ctactgtaac	3900
	ctctgaattc	gagagctaca	aagtccgagt	tcataatgtt	ctaaaaaac	agaaaaataa	3960
	atctatgtct	caaggctgaaa	ctgagggcgc	taaacaaag	agggaaatc	tggaaatgct	4020
	gattgaccag	ctaaaaatca	aattacaaga	tagccaaaat	aacttacaga	ttaatgtatc	4080
	tgaaacttcaa	acattgcagt	ctgaacatga	tacactgcta	gaaaggcaca	acaagatgct	4140
50	gcaggaaaact	gtgtccaaag	aggcggaact	ccgggaaaaa	ttgtgttcaa	tacagtcaga	4200
	gaacatgatg	atgaaatctg	aacatacaca	gactgtgagt	cagctaacat	cccagaacga	4260
	ggtcccttcga	aatagcttcc	gagatcaagt	gcgacatttg	caggaagaac	acagaaagac	4320
	agtggagaca	ttcagcagc	agctctccaa	gatggaagca	cagctcttcc	agcttaagaa	4380
	tgaaccgacc	acaagaagcc	cagtttccct	tcaacaatct	ttgaagaacc	ttcgagaag	4440
55	gagaaacaca	gacctccgc	ttctagacat	gcacactgta	acccgggaag	agggagaagg	4500
	catggagaca	actgatacgg	agtctgtgtc	ttccgcccgc	acatacacac	agtctttaga	4560
	gcagctgctt	aactctcccg	aaactaaact	tgagctcca	ttatggcatg	ctgaattttac	4620
	caaagaagaa	ttggttcaga	agctcagttc	caccacaaaa	agtgcagatc	acttaaacgg	4680
	cctgcttcgg	gaaacagaag	caaccaatgc	aattcttatg	gagcaaatga	agcttctcaa	4740
60	aagtgaataa	agaagattgg	aaaggaatca	agagcgagag	aagtctgcag	ctaacctgga	4800
	atacttgaag	aacgtcttgc	tgagttcat	tttcttgaaa	ccaggtagtg	aaagagagag	4860
	acttcttctc	gttataaata	cgatgttgca	gctcagccct	gaagaaaagg	gaaaacttgc	4920
	tgcggttgct	caaggtgagg	aagaaaatgc	ttcccgttct	tctggatggg	catcctatct	4980
	tcatagttgg	tctggacttc	gataggttga	tggaaggaat	atttttatta	accaaataga	5040
65	atctattttac	aaaaatggtt	cacgtatatt	accacaattc	ttttgtcaaa	aagtgtgtat	5100
	atatgtttgc	atctacatat	atttgtacat	ctatatgaca	gatgtatttt	aaaagtttca	5160
	tcttgaagta	aaagtacaac	agcttgaagt	gttgatagca	ggccacagcc	ctctaactca	5220

5	tgtgatttcc	catgcatgct	gccagaataa	aaccaccagg	aatgaattca	ctccccactt	5280
	ctctggaacc	tcaggaccgc	cccattttctc	ggcagtagctg	tgaatttttga	agttaaacta	5340
	aatttttgga	ccataccaac	tggaatttag	gcttttaaaaa	taatgttttca	aggccagggtg	5400
	tggtgattca	tgcctgaaat	cccactactt	tgggaggctg	aggctggaga	attgcttgag	5460
	gctagtggagc	tgtgactccc	actgcactcc	agctcgggga	acagagcgag	accttgtctc	5520
	taaaaaaat	agtaataaaa	taaaaaataac	gttttatgac	tatttattgc	aaggctcagag	5580
	ttacagattg	ttataaattg	ttgagaaatt	tttgtgatta	gaatatgaag	gaaaaagctt	5640
	tgttggtaaa	agtgcacatgt	taaggggcta	tgaagtaaat	atgctgcagt	taattgtgct	5700
	aagttaaaat	acagtttagt	tattttgcttt	aaaaataaact	cttctttttt	tctttaaagt	5760
10	atactatctc	aaaactcatt	atgtttgtcag	agccctagag	ctggctagtg	taacactgac	5820
	tatgagtagg	tgggcccacc	acttgagttg	aggtgatttc	atggtgtctt	tccaggctct	5880
	tgatagggtg	tactgcatg	caagccatga	atctgttttg	agaatcctct	ccattttccc	5940
	aaataaaaac	ctatcacaaac	agtgcata	tcactcagca	ttggatctaa	atataaaagt	6000
	ggtgcttttca	gtgttttttg	cagatagtgt	tcataaagct	ttccatcaga	agggatttta	6060
15	gacaccttag	aggtccgtgc	tacatcgtca	cagttcctcc	gaataacctt	aggtggtagt	6120
	gttacttgcc	tttgacacct	ctgcatatgt	tttaatgact	agatccaaac	tgtgttggtc	6180
	ttaaatcaaa	aattggataa	tttgtaatat	ttatgtgtta	atcacacagt	atgctctctg	6240
	aagtctctct	aagccttcag	tttatactct	taatttaatt	ttctttctga	gctggagaac	6300
	tggtcttgca	ctttgggttac	acagaacatt	ggtttccaat	tcagtttaac	tgaaatttgc	6360
20	tgctgatatg	ttgagtttgt	tcttttaaaa	atagctcata	tatctcatct	ttcctcctgt	6420
	cttagaagaa	cagacctaac	tagtgaatgt	attaatgaaa	atgcatctat	ttcagagctg	6480
	acatgaagag	tttagttttt	tacttttata	cactgtgaat	atgagtatgc	cagctgcata	6540
	cgatgtaact	aatcatatct	aaatatatct	cactttctct	ttgactttag	accttttgaa	6600
	gtctgtataa	acttgttttg	aaatatagtc	tctgcttacg	aatgtcataa	caaaataatt	6660
25	ttttgcatga	taaaaaatta	ctttgattac	aaaaggcgta	ttctttcatg	gtttctgcaa	6720
	tgagaggaag	tgtaatgatt	attttaatat	ttctattaaa	tatgtttaac	tgt	6773
30	<210> 28						
	<211> 2619						
	<212> DNA						
	<213> Homo sapiens						
35	<400> 28						
	atggccacag	cttgtaaaag	atcaggagaa	cctcagttctg	acgacattga	agctagccga	60
	atgaagcgag	cagctgcaaa	gcatctaata	gaacgctact	accaccagtt	aactgagggc	120
	tgtggaaatg	aagcctgcac	gaatgagttt	tgtgcttcct	gtccaaacttt	tcttcgtatg	180
	gataataatg	cagcagctat	taaaagccctc	gagctttata	agattaatgc	aaaactctgt	240
40	gatcctcatc	cctccaagaa	aggagcaagc	tcagcttacc	ttgagaactc	gaaaggtgcc	300
	cccaacaact	cctgctctga	gataaaaaatg	aacaagaaag	gcgctagaat	tgattttaaa	360
	gatgtgactt	acttaacaga	agagaaggta	tatgaaattc	ttgaattatg	tagagaaaga	420
	gaggattatt	cccctttaat	ccgtgtttat	ggaagagttt	tttctagtgc	tgaggcattg	480
	gtcacagagct	tccggaaggt	taaaacaacac	accaaggaag	aactgaaatc	tcttcaagca	540
45	aaagatgaag	acaaagatga	agatgaaaag	gaaaaagctg	catgttctgc	tgctgctatg	600
	gaagaagact	cagaagcatc	ttcctcaagg	ataggtgata	gctcacaggg	agacaacaat	660
	ttgcaaaaat	taggccttga	tgatgtgtct	gtggatattg	atgccattag	aagggtctac	720
	accagattgc	tctctaattga	aaaaattgaa	actgcctttc	tcaatgcaat	tgtatatttg	780
	tcacctaaac	tggaatgtga	cttgacgtat	cacaatgtat	actctcgaga	tcctaattat	840
50	ctgaatttgt	tcattatcgg	aatggagaat	agaaatctcc	acagtcctga	atatctggaa	900
	atggctttgc	cattattttg	caaagcgatg	agcaagctac	cccttgacgc	ccaaggaaaa	960
	ctgatcacag	tgtggtctaa	atacaatgca	gaccagattc	ggagaatgat	ggagacattt	1020
	cagcaactta	ttacttataa	agtcataatga	aatgaattta	acagtcgaaa	tctagtgaat	1080
	gatgatgatg	ccattgttgc	tgcttcgaag	tgctttgaaa	tggtttacta	tgcaaatgta	1140
55	gtgggagggg	aagtggacac	aaatcacaat	gaagaagatg	atgaagagcc	catccctgag	1200
	tccagcgagc	tgacacttca	ggaacttttg	ggagaagaaa	gaagaaacaa	gaaaggtcct	1260
	cgagtggaac	ccctggaaac	tgaacttggt	gttaaaaacc	tggattgtcg	aaaaccactt	1320
	atcccttttg	aagagttttat	taatgaacca	ctgaatgagc	ttctagaagt	ggataaagat	1380
	tatacttttt	tcaaagttaga	aaacagagaac	aaattctctt	ttatgacatg	tccctttata	1440
60	ttgaatgctg	tcacaaagaa	tttgggatta	tattatgaca	atagaattcg	catgtacagt	1500
	gaacgaagaa	tcactgttct	ctacagctta	gttcaaggac	agcagttgaa	tccatatttg	1560
	agactcaag	ttagacgtga	ccatatcaga	cagtagtcac	ttgtccggct	agagatgatc	1620
	gctatggaaa	atcctgcaga	cttgaagaag	cagttgtatg	tggaatttga	aggagaacaa	1680
	ggagttgatg	agggaggtgt	ttccaaagaa	ttttttcagc	tggttgtgga	ggaaatcttc	1740
65	aatccagata	ttggtatgtt	cacatacgat	gaatctacaa	aattgttttg	gtttaatcca	1800
	tcttcttttg	aaactgaggg	tcagtttact	ctgattggca	tagtactggg	tctggctatt	1860
	tacaataaact	gtatactgga	tgtacatttt	cccatgggtg	tctacaggaa	gctaattgggg	1920



	aaaaaaggaa	cttttcgtga	cttgggagac	tctcaccag	ttctatatca	gagtttaaaa	1980
	gatttattgg	agtatgaagg	gaatgtggaa	gatgacatga	tgatcacttt	ccagatatca	2040
	cagacagatc	tttttggtaa	cccaatgatg	tatgatctaa	aggaaaatgg	tgataaaatt	2100
5	ccaattacaa	atgaaaacag	gaaggaaatt	gtcaatcttt	attctgacta	cattctcaat	2160
	aaatcagtag	aaaaacagtt	caaggctttt	cggagagggt	ttcatatggt	gaccaatgaa	2220
	tctcccttaa	agtacttatt	cagaccagaa	gaaattgaat	tgcttatatg	tggaagccgg	2280
	aatctagatt	tccaagcact	agaagaaact	acagaatatg	acggtggcta	taccagggac	2340
	tctgttctga	ttagggagtt	ctgggaaatc	gttcattcat	ttacagatga	acagaaaaga	2400
	ctcttcttgc	agtttacaaac	gggcacagac	agagacacctg	tgaggaggact	aggaaaatta	2460
10	aagatgatta	tagccaaaaa	tggcccagac	acagaaaagg	tacctacatc	tcatacttgc	2520
	tttaattgtc	ttttacttcc	ggaatactca	agcaaaagaa	aacttaaaga	gagattgttg	2580
	aaggccatca	cgtatgccaa	aggatttggc	atgctgtaa			2619
	<210>	29					
15	<211>	4263					
	<212>	DNA					
	<213>	Homo sapiens					
	<400>	29					
20	ggcgcgtccc	ctctcctcag	cagtagctct	atggttttcag	ggcggcaacg	tgcagcgtcc	60
	ttaccttgag	cctgtgcagt	tgccctcacc	ccggaatcca	tagtcactgt	gacgaggcgg	120
	gaggacttgg	gcgacaggta	gcctcccagt	cccacacgct	gcgggtccgc	gcctggccaa	180
	gccacctcga	cctgtgaagt	tggggcggt	accagcaaac	tccccctgtg	cagccgcggt	240
	ttccaagggg	tcaggaaaccg	ctgtgtttgt	ttcgtccgcg	tagccagggc	gggtcgcgga	300
25	gtactgtgcc	tgacccgacg	gtggcaagtc	tgacgcgtca	gccagagacc	ggtgcccggt	360
	gtaggagtgc	cagcctgggc	tgtgagcggc	tgctgggtag	acagacttgc	tttctcttac	420
	agcatgtcat	ttccaaaatg	catcgtggtg	cttctgcctt	aagtcctata	ggaagacact	480
	gccgccacta	gaccggtgct	tatggtcgcc	actgttattc	tgactcaggt	cccgtgtcat	540
	tgagcatatg	tatgaaaatg	ccttaggagg	gaaccatgga	gaagtatgtg	agactgcaga	600
30	agattggaga	aggttcattt	ggaaaagctg	ttcttgttaa	atcgacagag	gatggcagac	660
	attatgtcat	caaggaaaatt	aacatctcaa	gaatgtctga	taaagaaagg	caagaatcaa	720
	ggagagaagt	tgctgtattg	gcaaacatga	agcatccaaa	tattgtccaa	tataaagaat	780
	catttgaaga	aaatggctct	ctctacatag	taatggatta	ctgtgaagga	ggtgatttgt	840
	ttaaaagaa	aaatgctcag	aaaggcgctc	tgtttcaaga	agaccagatt	ttggactggg	900
35	ttgtgcagat	atgtttggct	ctgaagcatg	tacatgatag	aaaaattctt	caccgagaca	960
	taaagtcaca	gaacataatt	ctaaccataa	atgggacagt	gcagcttggg	gattttggaa	1020
	ttgctcgagt	tcttaatagt	actgtagagc	tggtctgaac	ttgcatagga	actccatact	1080
	acttgtcacc	tgaatctcgt	gaaaacaagc	cttataacaa	taaaagtga	atttgggctt	1140
	tgggctgtgt	cctttatgag	ttgtgtacac	ttaaacatgc	atttgaagct	ggaacatga	1200
40	aaaacctggt	actgaagata	atctccggat	cttttctctc	agtgtctcca	cattactcct	1260
	atgatctccg	cagcttgcgt	tctcagttat	ttaaaagaaa	tcctagggat	agaccatcag	1320
	tcaactccat	attggagaaa	ggttttatag	ctaaacgaat	cgaaaagttt	ctctcccctc	1380
	agcttattgc	agaagaat	tgtctaaaaa	cactttcaaa	gtttggacca	cagcctctcc	1440
	caggtaaaa	accagcatca	ggacaaggtg	tcagttcttt	tgccctgtct	cagaaaaatca	1500
45	caaagcctgc	tgctaaatc	ggagtgcctt	taacataata	gaagtatgga	gataaaaagt	1560
	tacttgagaa	aaaaccaccc	ccaaaacata	aacaggccca	tcaaattccc	gtgaagaaaa	1620
	tgaattcttg	agaagaaaag	aagaaaatgt	ctgaggaagc	agcaaaaaaa	agaaggttgg	1680
	aatttattga	gaaagaaaag	aagcaaaaag	atcagattag	gttcctgaag	gctgagcaga	1740
	tgaagcggca	agagaagcag	cggttgggga	ggataaaatg	ggccagggaa	caaggatgga	1800
50	ggaatgtttt	aagggctggt	ggaagcgtg	aagtaaaagg	ttcctttttt	ggcattggag	1860
	gggctgtctc	tccatcaccg	tggtctctct	gaggccagta	tgaacattac	catgccattt	1920
	ttgaccaa	gcagcggcta	agagcagaag	ataatgaagc	aagatggaag	gggggaatct	1980
	atggtcagtg	gctcccagaa	aggcaaaaag	gacacttagc	tgtagagaga	gccaaccaag	2040
	tggaagaatt	cctacagcgt	aaacgagaag	ctatgcagaa	taaagcccga	gccgaaggac	2100
55	acgtggttta	tttggcaaga	ctgaggcaaa	taagactaca	aaattttaat	gagcgccaac	2160
	agattaaagc	caaacttcgt	ggtgagaata	aagaagctga	tggtaccaa	ggacaagaag	2220
	caactgaaga	gactgacatg	aggctcaaaa	agatggagtc	acttaaggcg	caaacaaatg	2280
	cacgtgtctg	tgtactaaaa	gaacagctgg	aggaaaaaag	aaaggaagct	tatgaaagag	2340
	aaaagaaaag	atgggaagaa	catttgggtg	cgaagggttaa	aagctcagat	gttctcttgc	2400
60	ctttggaact	tcttgaaaca	ggtggttctc	catcaaagca	gcaggtgaag	cctgtcattt	2460
	ctgtgacttc	agctttgaaa	gaagtggggc	tggttggaag	tttaactgat	acccagggaag	2520
	aagaaaatga	aaagagtaac	agtgtattt	caagtaagcg	agaaatcctg	cgtaggctaa	2580
	atgaaaatct	taaaagctca	gaggatgaaa	gcataactca	ggttcttgtg	ggttcttgtg	2640
	agaccgttgg	tcacaaagat	gagagagagt	atgagacaga	aaatgccatt	tcctctgac	2700
65	gcaagaagtg	ggagatggga	ggtcagcttg	tgattcctct	cgatgcagtg	acactggata	2760
	catccttctc	tgcaaccgaa	aaacatactg	tgggagaggt	tattaaatta	gatttctaatg	2820
	gctctccaa	aaaagtctgg	gggaaaaacc	ctacagattc	tgtgctgaag	atacttggag	2880



	aagctgaatt	acagctatag	acagaactac	tagaaaaacac	atctttttaa	agtgaagttt	2940
	atgctgaaga	ggagaactac	aaacccttac	ttactgaaga	agagaatctg	cagtgcattt	3000
	caaaaagaat	aaatccatca	gctactgttg	attctactga	aacgaaaagt	ccaaagttta	3060
5	ctgaggtgtc	tccacaaatg	tcagaaggaa	atgtggaaga	acctgatgat	ttggaacag	3120
	aagttctaca	agagccaagt	agcacacaca	cagatgggag	tttgccacct	gttcttaatg	3180
	atgtgtggac	tagagagaag	gaagcagcta	aggaaactga	gttggaagat	aaggttgctg	3240
	tgcagcagag	tgaagtttgt	gaagatagaa	ttccaggga	cgtggacca	tcctgtaagg	3300
	atcagagaga	tcctgcagta	gacgattctc	cgcagtctgg	ctgtgatgta	gagaagtcag	3360
	tacagccaga	atcgattttc	cagaaagtgg	ttcattctaa	ggacttgaac	ttagtccagg	3420
10	cagttcattg	ctcaccagaa	gaaccaattc	caattcgtac	tcactctgat	tctccacca	3480
	aaactaagag	caagaattcc	ttactgattg	gactttcaac	tggtctgttt	gatgcaaa	3540
	atccaaagat	gctgaggacc	tgctcacttc	cagatcttct	caagctgttc	agaaccctaa	3600
	tgagcgttcc	cactgtgggg	gacgttcac	aagacagtct	tgaatcgat	gagctggaag	3660
	atgaaccaat	taaagaaggg	ccttctgatt	ccgaagacac	tgtatttgaa	gaaactgaca	3720
15	cagattttaca	agagcttcag	gcctcaatgg	agcagctgct	tagggagcaa	ccaggtgacg	3780
	aatacagtga	ggaggaagag	tctgttttaa	aaagcagcga	tgtggagcag	acagcaagag	3840
	ggacagatgc	cccagacgag	gaggacaacc	ccagcagcga	aagcccctga	acgaggaatg	3900
	gcactcagat	aatagtgcg	ctgagaccac	tagtgaatgt	gaatatgaca	gtgtctttaa	3960
	ccatttagag	gaactaagac	ttcacttgga	gcaagaaatg	ggctttgaaa	agttctttga	4020
20	ggttttatgag	aaagtaaagg	ctattcatga	ggatgaagat	gaaaatattg	aaatttggtc	4080
	aacaatagtt	gagaatattt	tgggcaatga	gcaccagcat	ctctatgcc	agattctgca	4140
	tttagtcag	gcagatggag	cctatcagga	agataatgat	gaataatcct	caggacattc	4200
	tttaatagtc	aactgtaaga	acacatttga	acttggtctca	taataacaagc	ttcctgggaa	4260
25	ata						4263
	<210> 30						
	<211> 1756						
	<212> DNA						
	<213> Homo sapiens						
30	<400> 30						
	tcgggagcag	ccgcgaagat	gccgttgga	ctgacgcaga	gccgagtga	gaagatctgg	60
	gtgcccgtgg	accacagggc	ctcgttgccc	agatcctgtg	ggccaaagct	gaccaactcc	120
	cccaccgtca	tcgtcatggt	gggctcccc	gcccggggca	agacctacat	ctccaagaag	180
35	ctgactcgct	acctcaactg	gattggcgct	cccacaaaag	tgttcaacgt	cggggagtat	240
	cgccgggagg	ctgtgaagca	gtacagctcc	tacaacttct	tccgccccga	caatgaggaa	300
	gccatgaaag	tccggaagca	atgtgcctta	gctgccttga	gagatgtcaa	aagctacctg	360
	gcgaaagaag	ggggacaaat	tgcggttttc	gatgccacca	atactactag	agagaggaga	420
	cacatgatcc	ttcattttgc	caaagaaaat	gactttaaag	cgtttttcat	cgagtcggtg	480
40	tgcgacgacc	ctacagttgt	ggcctccaat	atcatggaag	ttaaaatctc	cagcccgat	540
	tacaaagact	gcaactcggc	agaagccatg	gacgacttca	tgaagaggat	cagttgctat	600
	gaagccagct	accagccctc	cgaccccgac	aaatgcgaca	gggacttgct	gctgatcaag	660
	gtgattgacg	tgggcccggg	gttctctggt	aaccgggtgc	aggaccacat	ccagagccgc	720
	atcgtgtact	actgatgaa	catccacgtg	cagccgcgta	ccatctacct	gtgccggcac	780
45	ggcgagaacg	agcaacaact	ccagggcgcg	atcgggggcg	actcaggcct	gtccagccgg	840
	ggcaagaagt	ttgccagtgc	tctgagcaag	ttcgtggagg	agcagaacct	gaaggacctg	900
	cgctgttgga	ccagccagct	gaagagcacc	atccagacgg	ccgaggcgct	gcggctgccc	960
	tacgagcagt	ggaaggcgct	caatgagatc	gacgcggcg	tctgtgagga	gctgacctac	1020
	gaggagatca	gggacaccta	cctgaggag	tatgcgctgc	gggagcagga	caagtactat	1080
50	taccgctacc	ccaccgggga	gtcctaccag	gacctgttcc	agcgcttggg	gccagtgtac	1140
	atggagctgg	agcggcagga	gaatgtgctg	gtcatctgcc	accaggccgt	cctgcgctgc	1200
	ctgcttgctc	acttctctga	taagagtgc	gaggagatgc	cctacctgaa	atgccctctt	1260
	cacaccgtcc	tgaactgcac	gcctgtcgct	tatggctgcc	gtgtggaatc	catctacctg	1320
	aacgtggagt	ccgtctgcac	acaccgggag	aggtcagagg	atgcaaagaa	gggacctaac	1380
55	ccgctcatga	gacgcaatag	tgtaaccccg	ctagccagcc	ccgaaccac	caaaaagcct	1440
	cgcatcaaca	gctttgagga	gcatgtggcc	tccacctcgg	ccgccctgcc	cagctgcctg	1500
	cccccgagg	tgccacgca	gctgcctgga	caaaacatga	aaggctcccg	gagcagcgct	1560
	gactcctcca	ggaaacactg	aggcagacgt	ctcggttcca	ttcattttcc	atttctgcag	1620
	cttagcttgt	gtcctgccct	ccgcccgagg	caaaacgtat	cctgaggact	tcttccggag	1680
60	aggggtgggt	ggagcagcgg	gggagccttg	gccgaagaga	accatgcttg	gcaccgtctg	1740
	gtccccctcg	gccgct					1756
	<210> 31						
	<211> 1661						
	<212> DNA						
65	<213> Homo sapiens						

<400> 31							
5	tgctgcagcc	gctgccgccc	attccggatc	tcattgccac	gcgcccccca	cgaccgccc	60
	acgtgcattc	ccgattccct	ttggttccaa	gtccaatatg	gcaactctaa	aggatcagct	120
	gattttataa	cttctaaaagg	aagaacagac	ccccagaat	aagattacag	ttgttgggg	180
	tggtgctgtt	ggcatggcct	gtgccatcag	tatcttaatg	aaggacttgg	cagatgaact	240
	tgctcttgtt	gatgtcatcg	aagacaaatt	gaaggagag	atgatggatc	tccaacatgg	300
10	cagccttttc	cttagaacac	caaagattgt	ctctggcaaa	gactataatg	taactgcaaa	360
	ctccaagctg	gtcattatca	cggtctgggc	acgtcagcaa	gaggagaaa	gccgtcttaa	420
	tttgggtccag	cgtaacgtga	acataatttaa	attcatcatt	cctaattgtt	taaaatacac	480
	cccgaactgc	aagttgctta	ttgtttccaa	tccagtggat	atcttgacct	acgtggcctt	540
	gaagataagt	ggttttccca	aaaaccgtgt	tattggaagt	ggttgcaatc	tggattcagc	600
15	ccgattccgt	tacctgatgg	gggaaaggct	gggagttcac	ccattaagct	gtcatgggtg	660
	ggtccttggg	gaacatggag	attccagtgt	gcctgtatgg	agtggatga	atgttgctgg	720
	tgtctctctg	aagactctgc	accagatttt	agggactgat	aaagataagg	aacagtggaa	780
	agaggttcac	aagcaggtgg	ttgagagtgc	ttatgaggtg	atcaaaactca	aaggctacac	840
	atcctgggct	attggactct	ctgtagcaga	tttggcagag	agtataatga	agaatcttag	900
20	gcgggtgcac	ccagtttcca	ccatgattaa	gggtctttac	ggaataaagg	atgatgtctt	960
	ccttagtggt	ccttgcatct	tgggacagaa	tggaaatctca	gaccttgtga	aggtgactct	1020
	gacttctgag	gaagaggccc	gtttgaagaa	gagtgcagat	acactttggg	ggatccaaaa	1080
	ggagctgcaa	ttttaaagtc	ttctgatgtc	atatcatttc	actgtctagg	ctacaacagg	1140
	attctaggtg	gaggttgtgc	atgttgcctt	ttttatctga	tctgtgatta	aagcagtaat	1200
25	atcttaagat	ggactgggaa	aaacatcaac	tctgaagtt	agaaataaga	atggtttgta	1260
	aaatccacag	ctatatctct	atgctggatg	gtattaatct	tgtgtagtct	tcaactgggt	1320
	agtgtgaaat	agttctgcca	cctctgacgc	accactgcca	atgctgtacg	tactgcattt	1380
	gcccccttag	ccaggtggat	gtttaccgtg	tggtatataa	cttcctggct	ccttcactga	1440
	acatgcctag	tccaacattt	tttccagtg	agtcacatcc	tgggatccag	tgtataaatc	1500
30	caatatcatg	tcttgtgcat	aattcttcca	aaggatctta	ttttgtgaac	tatatcagta	1560
	gtgtacatta	ccatataatg	taaaaagatc	tacatacaaa	caatgcaacc	aactatccaa	1620
	gtgttatacc	aactaaaacc	cccaataaac	cttgaacagt	g		1661
<210> 32							
<211> 4169							
<212> DNA							
<213> Homo sapiens							
<400> 32							
40	ggcggcttcc	aggtgggccc	gcaaggccgt	ggctctgctt	tgtgcctctg	acctgctgct	60
	gctgctgcta	ctgctaccac	cgccctgggtc	ctgcgcccgc	gaaggctcgc	ccgggacgcc	120
	cgacagatct	acccacctc	cccgaagaa	gaagaaggat	attcgcgatt	acaatgatgc	180
	agacatggcg	cgctctctgg	agcaatggga	gaaagatgat	gacattgaag	aaggagatct	240
	tccagagcac	aagagacctt	cagcacctgt	cgacttctca	aagatagacc	caagcaagcc	300
45	tgaagacata	ttgaaaatga	cgaaaaaagg	gaagactctc	atgatgtttg	tactgtatc	360
	aggaagccct	actgagaagg	agacagagga	aattacgagc	ctctggcagg	gcagcctttt	420
	caatgccaac	tatgacgtcc	agaggttcat	tgtgggatca	gaccgtgcta	tcttcatgct	480
	tcgcgatggg	agctacgcct	gggagatcaa	ggactttttg	gtcgggtcaag	acaggtgtgc	540
	tgatgtaaat	ctggagggcc	aggtgtaccc	cgccaaagga	ggaggaagca	aagagaaaaa	600
50	taaaacaaag	caagacaagg	gcaaaaaaaa	gaaggaagga	gatctgaaat	ctcgtctctc	660
	caaggaagaa	aatcgagctg	ggaataaaaag	agaagacctg	tgatggggca	gcagtgcgcg	720
	gctgtggggg	gacaggtgga	cgtggagagc	tctttgcccc	gctcctgggg	tgggagtggg	780
	ctcaggcaac	tgacacccgg	atgacattct	agtgtcttct	agaaagggtc	tgccacatga	840
	ccagtttgtg	gtcaaaagaat	tactgcttaa	taggcttcaa	gtaagaagac	agatgttttc	900
55	taattaatac	tggaactga	caaattcatg	tttactataa	aatctcctta	catggaaatg	960
	tgactgtggt	gctttttccc	atttacactt	ggtgagtcac	caactctaact	gagattccac	1020
	tcccctccaa	gcacctgctg	tgattgggtg	gcctgctctg	atcagatagc	aaattctgat	1080
	cagagaagac	tttaaaactc	ttgacttaat	tgagtaaaact	cttcacgcca	tatacatcat	1140
	tttcattatg	ttaaaggtaa	aatatgcttt	gtgaactcag	atgtctgtag	ccaggaagcc	1200
60	aggggtgtga	aatccaaaat	ctatgcagga	aatgcggaga	atagaaaata	tgtcacttga	1260
	aatcttaagt	agttttgaat	ttctttgact	tgaactttac	tcatcagtaa	gagaactctt	1320
	ggtgtctgtc	aggttttatg	tggtctgtaa	agttaggggt	tctgttttgt	ttccttattt	1380
	aggaagagat	actgctgggtg	tcgaggggtt	atatgttcca	tttaattgtga	cagttttaaa	1440
	ggattttaag	agggaaatcag	agtcctttgc	agagtggtgac	agacgactca	ataacctcat	1500
65	ttgtttctaa	acatttttct	ttgataaagt	ccgtaaatct	gtgctttcgt	atagagtaac	1560
	atgatgtgct	actgttgatg	tctgattttg	cggttcatgt	tagagcctac	tgtgaataag	1620
	agttagaaca	tttatataca	gatgtcattt	ctaagaacta	aaattctttg	ggaaaaacc	1680
	tcaattgtga	ttttaataaa	ttaaaagtag	cacattacat	gggttagaaa	tgtcagtgtt	1740
	aaagaatggg	acaaagtga	aagtgtatcc	ctctcttgcc	gccggtggta	gcttgtccca	1800
	gtggaagctg	ctgttaacaa	tttgtgcccc	cacatcccc	tccctgcccc	tccacccaaa	1860

	aaaagtacat	ttacttatgt	aaatgtactt	atggtgatgt	atgtttgttt	tggcctcaca	1920
	gcatctgttt	ccccttaatt	tggtagctgc	tcacatttcc	ctcgaaagaa	ccacaccctc	1980
	tgcattctca	gttcttttgc	ttggatggga	cattttgccct	gcagtcctcc	caccctccag	2040
5	gccatgccc	ctccagggtg	aggcctgtgt	gatctaccgt	actagggtac	taggccctga	2100
	aagaggcttt	tcttgttcct	cctgcatctt	gaacctggag	cgggagctgt	tgtaggcccc	2160
	gcccttggag	aagagaactg	tctgacagt	gggagagagc	gccacaccct	ggtggcataa	2220
	acgagtcctt	gaatcatgcc	gtggctgaac	caagccctgt	ctgtgggctt	tttctgttgt	2280
	actcagggca	gtttgatggg	gttactgtcc	tgcatagcc	taatggccca	gtataaagca	2340
10	gctgttttga	tgagataatt	gctttaatta	agcaaaaggt	agcaaaagctt	tcactccgcc	2400
	ctgtaccttc	tgtttccact	taggagcctt	cccattgtcag	aatgtgcaga	tctgtctcat	2460
	tgtttctgt	gcagtggtgc	cccacttcac	ccagtagttt	ctgtgtgtct	gttatgtact	2520
	aggtagctaca	aggtagccagg	acggtgtaga	tacagcctct	gctatcgtaa	aactcaatga	2580
	ttcgttgggg	gaagacaaat	gtcagtaatt	tacaaagtaa	aatggcagct	gttagaagta	2640
15	tgaaaggggc	agggtagggg	gaggtagaat	cttccctgag	caggttaaga	aaaccagagg	2700
	ccttctctga	gggcaagagg	aggagaggag	aaatagagta	aggcaggcag	aggaaccagt	2760
	ctgagctaag	accctgtggc	tagaagtggc	agagggagag	gcagcaggaa	ggccagcggg	2820
	gaggctgggg	cccagtgag	gcccaggttg	gaggagcgta	gcacatggag	tttggtagga	2880
	gtttggggac	ccctgggtga	tcttaattgt	gatgggtggg	gtgtgaaagg	cagtccaggt	2940
20	tgactgtgtt	gcacaggaga	agtgtatcaga	agaggacccc	agcaggtgtg	agccgtgagc	3000
	tgggagggtgc	ttcagtagtg	caggccatag	ctgaaggtgt	cctacatcag	cagggtgtatg	3060
	gtgaggtttg	aaccactgtt	tcactgcata	gtccctgctg	atggacactt	gagtgttcag	3120
	attttttgct	ggtatattca	gtgctgcagt	ggacattttc	atacaaaata	tttcggtaca	3180
	cttttgttta	tatctgaaag	gtaaaattcct	agcagtagaa	ttattagagc	aaacggaatt	3240
25	taacattttg	gtgtgtattg	ccaaattgcc	ctcccaagt	gtttagttag	cttacccttg	3300
	ccaacaatag	atctatcctt	gccagccttg	ggcatcacat	ttaccagttt	aatagattgt	3360
	aaaaccatat	cttaattggc	taccctgaag	ccaccatact	ggagaggctg	cgtacagtgt	3420
	ttcacgtaga	gagagggata	cccaggaggc	ccactgtctc	caaccccgag	tgcatgagtc	3480
	ttcccagccc	aggcacagac	atgtggataa	gatttaaaca	ttccagccc	cagccttcaa	3540
30	gcaatcctag	ttgacactga	ggggagccaa	cataagctga	gctgagaaac	agtctgcccc	3600
	gtctgcagat	tcattgagcaa	aagaaatgtt	gggtctgggt	cagtggctca	cgctgtaat	3660
	cccagtactt	tgggaggccg	agggtgggtg	atcagtttag	gtcaggagtt	tgagaccagc	3720
	ctggccaaca	tggtgaagcc	ctgtctctac	taaaaattag	ccgagtgtgg	tgggtgcggc	3780
	ctgtaatccc	agctactcag	gtggctgagg	caggagaatg	gcttgaaccc	gggaggcgga	3840
	ggttgcagtg	agccaagatc	aggccactgc	actccagcct	ggatgacggg	atgagactct	3900
35	gtctcaaaaa	aacgaaacaa	aaatttttta	agagaaatgt	catttgtttt	tgtttttgag	3960
	acagggtctc	actctgttgc	cctcactaga	gtgcagtagg	gatcacggct	cactgaagtc	4020
	tctacctacc	ggctcaattg	atcttcccac	cacagcctcc	caaatagctg	ggagaaatgt	4080
	cctgttttta	atgaatttgt	cttccttttt	gtcttgtttg	ttttaatatc	tagtgatcta	4140
40	ataaatttgg	atgatattct	ttgactatc				4169
	<210>	33					
	<211>	859					
	<212>	DNA					
	<213>	Homo sapiens					
45	<400>	33					
	catgccatgc	agcttacc	ccatcgagtc	tgggactatg	ctggagataa	ctatgttcat	60
	cgactggttg	caagtaaaac	agatggaaaa	atagtagagt	atgaatgtga	gggggatact	120
	tgccagggaag	agaaaaataga	tgctttacag	ttagagtatt	catatttact	aacaagccag	180
50	ctggaatctc	agcgaatcta	ctgggaaaaac	aagatagtct	ggatagagaa	ggacacagca	240
	gaggaaatta	acaacatgaa	gaccaagtgt	aaagaaacaa	ttgagaagt	tgataatcta	300
	gagcacaac	taaatgatct	cctaaaagaa	aagcagctct	tggaaagaaa	gtgcactcag	360
	ctaaacacaa	aagtggccaa	actcaccaac	gagctcaaa	aggagcagga	aatgaacaag	420
	tgtttgcgag	ccaaccaagt	cctcctgcag	aacaagctaa	aagaggagga	gagggtgtct	480
55	aaggagacct	gtgacaaaa	agatctgcag	atcacccaga	tccaggagca	gctgctgtgac	540
	gtcatgttct	acctggagac	acagcagaag	atcaaccatc	tgctgcccga	gacccggcag	600
	gaaatccagg	aggagacagat	caacatcgcc	atggcctcgg	cctcgagccc	tgctcttctg	660
	gggggcagtg	ggaagtggcc	ctccaggaa	ggcgagca	agaggggcaa	gtgaccttca	720
	gagcaacaga	tacccctgag	actgttctcc	ctgacactgt	gagagtgtgc	tgggaccttc	780
60	agctaaatgt	gagggtgggc	cctaataagt	acaagttagg	acgaaggccg	gccttcgtgg	840
	ccttagagat	ggatgaggc					859
	<210>	34					
	<211>	1070					
	<212>	DNA					
65	<213>	Homo sapiens					

<400> 34		
	gcgattgctg gggctgcagc gctgcctccg agaccgagag tgggtggagc gggctcttcc	60
	ggaagggtgc gataaggccg ggcgaggtgc ctgggatgct tctccccctc cgcgaggaag	120
	agatctaatt gggtagggcg ggtgtagact agcctgccga gccgcccgtc ggcacctgca	180
5	gcctcctggg cgcccgcggg cccggcgaga aagtgtgtaa agggagcgag gtggttggtc	240
	ctgggggtccg aggcgcgcct ctacgcctc gcccaacaga agccgcagtc ccgtgggggtc	300
	tggagacgca gtttccttgt taatgacaat aaatccctgc tccccctgcc tcagacatct	360
	acgcagcgaa atcgagcctg gccttgaggg tccacaccgc gaggaagatg cgtgcgcca	420
	ttccagagcc taagcctgga gacctgattg agattttctc ccctttctac agacactggg	480
10	ccatctatgt tggcgatgga tatgtggttc atctggcccc tccaagttag gtcgcaggag	540
	ctggtgcagc cagtgtcatg tccgccctga ctgacaaggc catcgtgaag aaggaattgc	600
	tgtatgatgt ggccgggagt gacaagtacc aggtcaacaa caaacatgat gacaagtact	660
	cgccgctgcc ctgacgaaa atcatccagc ggccggagga gctggtgggg caggagggtgc	720
	tctacaagct gaccagttag aactgcgagc actttgtgaa tgagctgcgc tatggagtgc	780
15	cccgcagtga ccaggtcaga gatgtcatca tcgctgcaag cgttgccagga atgggcttgg	840
	cagccatgag ccttatttga gtcatgttct caagaaacaa gcgacaaaag caataactga	900
	aaaagactgt ctgtcagcga tgactttata catcaagggg gtcttgtttt gctagagagt	960
	ttgggggttg gtttgtggat ttcatgttga ttataataa ggcttatttt cacagaataa	1020
20	aataaagcaa aacgaggag gattttattg ggggagtga gccccaaaaa	1070
<210> 35		
<211> 460		
<212> DNA		
<213> Homo sapiens		
<400> 35		
	cttttctctc catgtcgcca ccgaggtgcc acgctgaga cttctccgcc gcctccgcc	60
	cagacgccgc cgcgatgcgc tacgtgcct cctacctgct ggctgcccta gggggcaact	120
	cctccccag cgccaaggac atcaagaaga tcttgacag cgtgggtatc gaggcggacg	180
30	acgacgggct caacaagggt atcagttagc tgaatggaaa aaacattgaa gacgtcattg	240
	cccagggtat tggcaagctt gccagtgtac ctgctgggtg ggctgtagcc gtctctgtcg	300
	ccccaggctc tgcagccctt gctgctggtt ctgccctgc tgcagcagag gagaagaaag	360
	atgagaagaa ggagaggtct gaagagtcag atgatgacat gggatttggc ctttttgatt	420
35	aaattcctgc tcccctgcaa ataaagcctt tttacacatc	460
<210> 36		
<211> 4416		
<212> DNA		
<213> Homo sapiens		
<400> 36		
	gccgcacagg gttttatagg atcacattga caaaagtacc atggagtgtt atgagtcagc	60
	atattttatt gttcttattc cttcaatagt tattacagta attttcctct tcttctggct	120
	tttcatgaaa gtaaacattat atgatgaagt tcttgcaaaa cagaaaagag aacaaaagct	180
45	tattcctacc aaaacagata aaaagaaagc agaaaagaaa aagaataaaa agaaagaaat	240
	ccagaatgga aacctccatg aatccgactc tgagagtgtt cctcgagact ttaaattatc	300
	agatgctttg gcagtagaag atgatcaagt tgcacctgtt ccattgaatg tcgttgaaac	360
	ttcaagttagt gttagggaaa gaaaaaagaa gaaaaagaaa caaaagcctg tgcttgaaag	420
	gcaggtcatc aaagaaaagt acgcatcaaa gattcctggc aaaaaagtag aacctgtccc	480
50	agttactaaa cagcccaccc ctccctctga agcagctgcc tcgaagaaga aaccagggca	540
	gaagaagtct aaaaatggaa gcgatgacca ggataaaaag gtggaactc tcatggatcc	600
	atcaaaaagg caagaagcat tgcccctcca ccaagagact aaacaagaaa gtggatcagg	660
	gaagaagaaa gcttcatcaa agaaacaaaa gacagaaaat gtcttcgtag atgaaccct	720
	tattcatgca actacttata ttcccttgat ggataatgct gactcaagtc ctgtggtaga	780
55	taagagagag gttattgatt tgcttaaacc tgaccaagta gaaggatcc agaaatctgg	840
	gactaaaaaa ctgaagaccg aaactgacaa agaaaatgct gaagtgaagt ttaaagattt	900
	tcttctgtcc ttgaagacta tgatgttttc tgaagatgag gctctttgtg ttgtagactt	960
	gctaaggag aagtctggtg taatacaaga tgctttaaag aagtcaagta agggagaatt	1020
	gactacgctt atacatcagc ttcaagaaaa ggacaagtta ctgctgctg tgaaggaaag	1080
60	tgctgctgct acaaaggatc ggtgtaagca gtttaaccag gaaatgatga cagagaaaga	1140
	aagaagcaat gtggttataa caaggatgaa agatcgaatt ggaacattag aaaaggaaca	1200
	taatgtattt caaaacaaaa tacatgtcag ttatcaagag actcaacaga tgcagatgaa	1260
	gtttcagcaa gttcgtgagc agatggaggc agagatagct cacttgaagc aggaaaatgg	1320
	tatactgaga gatgcagtca gcaacactac aaatcaactg gaaagcaagc agtctgcaga	1380
65	actaaataaa ctacgccagg attatgctag gttgtgtaag gagctgactg agaaaacagg	1440
	aaagctacag caagaggaag tccaaaagaa gaatgctgag caagcagcta ctcagttgaa	1500
	ggttcaacta caagaagctg agagaaggtg ggaagaagtt cagagctaca tcaggaagag	1560

	aacagcggaa	catgaggcag	cacagcaaga	tttacagagt	aaattttgtgg	ccaaagaaaa	1620
	tgaagtacag	agtctgcata	gtaagcttac	agataccttg	gtatcaaaac	aacagttgga	1680
	gcaaagacta	atgcagttaa	tggaatcaga	gcagaaaagg	gtgaacaaag	aagagtctct	1740
5	acaaatgcag	gttcaggata	ttttggagca	gaatgaggct	ttgaaagctc	aaattcagca	1800
	gttccattcc	cagatagcag	cccagacctc	cgcttcagtt	ctagcagaag	aattacataa	1860
	agtgattgca	gaaaaggata	agcagataaa	acagactgaa	gattcttttag	caagtgaacg	1920
	tgatcgttta	acaagtaaag	aagaggaaact	taaggatata	cagaatatga	atttcttatt	1980
	aaaagctgaa	gtgcagaaat	tacaggccct	ggcaaatgag	caggctgctg	ctgcacatga	2040
10	attggagaag	atgcaacaaa	gtgtttatgt	taaagatgat	aaaataagat	tgctggaaga	2100
	gcaactacaa	catgaaattt	caaacaaaaat	ggaagaattt	aagattctaa	atgacaaaaa	2160
	caaagcatta	aaatcagaag	ttcagaagct	acagactctt	gtttctgaac	agcctaataa	2220
	ggatgttgtg	gaacaaatgg	aaaaatgcat	tcaagaaaaa	gatgagaagt	taaagactgt	2280
	ggaagaatta	cttgaactg	gacttattca	ggtggcaact	aaagaagagg	agctgaatgc	2340
15	aataagaaca	gaaaattcat	ctctgacaaa	agaagttaa	gacttaaaag	ctaagcaaaa	2400
	tgatcaggtt	tcttttgctt	ctctagttag	agaacttaag	aaagtgatcc	atgagaaga	2460
	tggaaaagatc	aagtctgtag	aagagcttct	ggaggcagaa	cttctcaaag	ttgctaacia	2520
	ggagaaaact	gttcaggatt	tgaacacagga	aataaaggct	ctaaaagaag	aaataggaaa	2580
	tgtccagctt	gaaaaggctc	aacagttatc	tatcacttcc	aaagttcagg	agcttcagaa	2640
20	cttattaaaa	ggaaaaggagg	aacagatgaa	taccatgaag	gctgttttgg	aagagaaga	2700
	gaaagaccta	gccaatcacg	ggaagtgggt	acaggatctt	caagaagaaa	atgaatcttt	2760
	aaaagcacat	gttcagggaag	tagcacaaca	taacttgaaa	gaggcctctt	ctgcatcaca	2820
	gtttgaagaa	cttgagattg	tggtgaaaga	aaagggaaat	gaattgaaga	ggttagaagc	2880
	cattgctaata	gagagggaga	gtgatctttc	tagcaaaaaca	cagctgttac	aggatgtaca	2940
25	agatgaaaac	aaattgttta	agtcccaaat	tgagcagctt	aaacaacaaa	actaccaaca	3000
	ggcatcttct	tttccccctc	atgaagaatt	attaaaagta	atttcagaaa	gagagaaaga	3060
	aataagtggt	ctctggaatg	agtttagattc	tttgaaggat	gcagttgaac	accagaggaa	3120
	gaaaaacaat	gaaaggcagc	aacaggtgga	agctgttgag	ttggaggcta	aagaagttct	3180
	caaaaaatta	tttccaaagg	tgtctgtccc	ttctaatttg	agttatggtg	aatgggtgca	3240
30	tggatttgaa	aaaaaggcaa	aagaatgtat	ggctggaact	tcagggtcag	aggaggttaa	3300
	ggttctagag	cacaagttga	aagaagctga	tgaatgcac	acattgttac	agctagagtg	3360
	tgaaaaatata	aaatccgtcc	ttgcagaaac	agaaggaatt	ttacagaagc	tacagagaag	3420
	tgttgagcaa	gaagaaaata	aatggaaagt	taaggtcgat	gaatcacaca	agactattaa	3480
	acagatgcag	tcatcattta	catcttcaga	acaagagcta	gagcgattaa	gaagcgaaaa	3540
35	taaggatatt	gaaaatctga	gaagagaacg	agaacatttg	gaaatggaac	tagaaaaggc	3600
	agagatggaa	cgatctacct	atgttacaga	agtcagagag	ttgaaggcac	agttaaatga	3660
	aacactcaca	aaacttagaa	ctgaacaaaa	tgaagacag	aaggtagctg	gtgatttgca	3720
	taaggctcaa	cagtcactgg	agcttatcca	gtcaaaaata	gtaaaagctg	ctggagacac	3780
	tactgttatt	gaaaatagtg	atgtttcccc	agaaacggag	tcttctgaga	aggagacaat	3840
	gtctgtaagt	ctaaatcaga	ctgtaacaca	gttacagcag	ttgcttcagg	cggtaaacca	3900
40	acagctcaca	aaggagaaag	agcactacca	ggtgttagag	tgaagtaatt	gggaaactgt	3960
	tcatttgagg	ataaaaaagg	cattgtatta	tattttgcca	aattaaagcc	ttatttatgt	4020
	tttcaccctt	tctactttgt	cagaaacact	gaacagagtt	ttgtcttttc	taatccttgt	4080
	tagactactg	attttaaagaa	ggaaaaaaa	aagccaactc	tgtagacacc	ttcagagttt	4140
45	agttttataa	taaaaaactgt	ttgaataatt	agacctttac	attcctgaag	ataaacatgt	4200
	aatcttttat	cttattttgc	tcaataaaat	tgttcagaag	atcaaagtgg	ttaaagacaat	4260
	gtaaaattta	acattttta	actgatgttg	tacactgttt	tacttaacat	tttgggaagt	4320
	aactgcctct	gacttcaact	caagaaaaca	cttttttgtt	gctaattgtaa	tcggtttttg	4380
	taatggcgctc	agcaataaaa	aggatgctta	ttatttc			4416
50		<210> 37					
		<211> 628					
		<212> DNA					
		<213> Homo sapiens					
55		<220>					
		<221> unsure					
		<222> 284..284					
		<223> n = a, c, g or t					
60		<220>					
		<221> unsure					
		<222> 594..594					
		<223> n = a, c, g or t					
65		<220>					
		<221> unsure					
		<222> 597..597					

<223> n = a, c, g or t  
 <220>  
 <221> unsure  
 5 <222> 607..607  
 <223> n = a, c, g or t  
 <400> 37  
 10 ggcacgagaa gaagtctaaa aatggaagcg atgaccaggg ataaaaaggt ggaaactctc 60  
 atggtaccat caaaaaggca agaagcattg cccctccacc aagagactaa acaagaaagt 120  
 ggatcaggga agaagaaagc ttcacaaag aaacaaaaga cagaaaatgt cttcgtagat 180  
 gaaccctta ttcacgaac tacttatatt cctttgatgg ataatgctga ctcaagtcct 240  
 gtggtagata agagagagggt tatttatattg cttaaacctg accnagtaga agggatccag 300  
 15 aaatctggga ctaaaaaact gaagacccaa actgacaaaag aaaatgctga agtgaagttt 360  
 aaagattttc ttctgtcctt taagactatg atgttttctg aagatgaggc tctttgtgtt 420  
 gtagacttgc taaaggagaa gtctcgtgta atacaagatg ctttaaagaa gtcaagtaag 480  
 ggagaattga ctacgcttat acatcagctt caagaaaagg accaagttac tcgctgctgt 540  
 gaaggagat gctgctgcta caaaggatcc gtgtaagcag ttacccagc aatnatncca 600  
 20 aagaaanaca gagcaatttg gtatacaa 628  
 <210> 38  
 <211> 730  
 <212> DNA  
 25 <213> Homo sapiens  
 <220>  
 <221> unsure  
 <222> 127..127  
 30 <223> n = a, c, g or t  
 <220>  
 <221> unsure  
 <222> 260..260  
 35 <223> n = a, c, g or t  
 <220>  
 <221> unsure  
 <222> 307..307  
 40 <223> n = a, c, g or t  
 <220>  
 <221> unsure  
 <222> 329..329  
 45 <223> n = a, c, g or t  
 <220>  
 <221> unsure  
 <222> 336..336  
 50 <223> n = a, c, g or t  
 <220>  
 <221> unsure  
 <222> 473..473  
 55 <223> n = a, c, g or t  
 <220>  
 <221> unsure  
 <222> 487..487  
 60 <223> n = a, c, g or t  
 <220>  
 <221> unsure  
 <222> 494..494  
 65 <223> n = a, c, g or t  
 <220>  
 <221> unsure

```

5      <222> 563..563
      <223> n = a, c, g or t

      <220>
      <221> unsure
      <222> 640..640
      <223> n = a, c, g or t

10     <220>
      <221> unsure
      <222> 681..681
      <223> n = a, c, g or t

15     <220>
      <221> unsure
      <222> 689..689
      <223> n = a, c, g or t

20     <220>
      <221> unsure
      <222> 708..708
      <223> n = a, c, g or t

25     <220>
      <221> unsure
      <222> 714..714
      <223> n = a, c, g or t

30     <220>
      <221> unsure
      <222> 728..728
      <223> n = a, c, g or t

35     <220>
      <221> unsure
      <222> 730..730
      <223> n = a, c, g or t

40     <400> 38
      aaagaattcg gcacgagtga agaccgaaac tgacaaagaa aatgctgaag tgaagttaa 60
      agattyyctt ctgtccttga agactatgat gttttctgaa gatgaggctc tttgtgttgt 120
      agactyncta aaggagaagt ctggtgtaat acaagatgct ttaaagaagt caagtaaggg 180
      agaattgact acgcttatac atcagcttca agaaaaggac aagtyactcg ctgctgtgaa 240
      ggaagatgct gctgcyacan aggatcggtg taagcagtta acccaggaaa tgatgacaga 300
45     gaaaganaga agcaatgtgg ttataacang gatganagay cgaatyggaa cattagaaaa 360
      ggracataat gtatyycaaa acaaaaataca tgtcagtyat caagagacyc aacagatgca 420
      gatgaagttt cagcaagttc gtgagcagat ggaggcagag atagcycact tgnagcagga 480
      aaatgggnata ctgngagaat gcagtcagca acactacaaa tcaactggaa agcaagcagt 540
      ctgcagaact aaataaaacta cgncaggatt atgctaggtt ggtgaatgag cctgactgag 600
50     aaacaggaaa gctacagcaa gaggaagtca aagaagaatn ctgagcaagc agctactcag 660
      ttgaaggttc caactacacg nacgcctgng agaaggtggg ggaggcgntt cagngcctac 720
      atcaggngng                                     730

55     <210> 39
      <211> 571
      <212> DNA
      <213> Homo sapiens

60     <220>
      <221> unsure
      <222> 37..37
      <223> n = a, c, g or t

65     <220>
      <221> unsure
      <222> 42..42
      <223> n = a, c, g or t

```

5       <220>  
      <221> unsure  
      <222> 64..64  
      <223> n = a, c, g or t

10       <220>  
      <221> unsure  
      <222> 67..67  
      <223> n = a, c, g or t

15       <220>  
      <221> unsure  
      <222> 73..73  
      <223> n = a, c, g or t

20       <220>  
      <221> unsure  
      <222> 86..99  
      <223> n = a, c, g or t

25       <220>  
      <221> unsure  
      <222> 106..107  
      <223> n = a, c, g or t

30       <220>  
      <221> unsure  
      <222> 117..117  
      <223> n = a, c, g or t

35       <220>  
      <221> unsure  
      <222> 139..139  
      <223> n = a, c, g or t

40       <220>  
      <221> unsure  
      <222> 153..153  
      <223> n = a, c, g or t

45       <220>  
      <221> unsure  
      <222> 243..243  
      <223> n = a, c, g or t

50       <220>  
      <221> unsure  
      <222> 257..257  
      <223> n = a, c, g or t

55       <220>  
      <221> unsure  
      <222> 264..264  
      <223> n = a, c, g or t

60       <220>  
      <221> unsure  
      <222> 319..319  
      <223> n = a, c, g or t

65       <220>  
      <221> unsure  
      <222> 337..337  
      <223> n = a, c, g or t

      <220>



<221> unsure  
<222> 341..341  
<223> n = a, c, g or t

5       <220>  
       <221> unsure  
       <222> 353..353  
       <223> n = a, c, g or t

10       <220>  
       <221> unsure  
       <222> 362..362  
       <223> n = a, c, g or t

15       <220>  
       <221> unsure  
       <222> 373..373  
       <223> n = a, c, g or t

20       <220>  
       <221> unsure  
       <222> 379..379  
       <223> n = a, c, g or t

25       <220>  
       <221> unsure  
       <222> 384..385  
       <223> n = a, c, g or t

30       <220>  
       <221> unsure  
       <222> 419..419  
       <223> n = a, c, g or t

35       <220>  
       <221> unsure  
       <222> 421..421  
       <223> n = a, c, g or t

40       <220>  
       <221> unsure  
       <222> 427..427  
       <223> n = a, c, g or t

45       <220>  
       <221> unsure  
       <222> 433..433  
       <223> n = a, c, g or t

50       <220>  
       <221> unsure  
       <222> 440..442  
       <223> n = a, c, g or t

55       <220>  
       <221> unsure  
       <222> 444..444  
       <223> n = a, c, g or t

60       <220>  
       <221> unsure  
       <222> 448..448  
       <223> n = a, c, g or t

65       <220>  
       <221> unsure  
       <222> 450..450

5           <223> n = a, c, g or t  
          <220>  
          <221> unsure  
          <222> 455..455  
          <223> n = a, c, g or t  
  
10           <220>  
          <221> unsure  
          <222> 486..486  
          <223> n = a, c, g or t  
  
15           <220>  
          <221> unsure  
          <222> 488..489  
          <223> n = a, c, g or t  
  
20           <220>  
          <221> unsure  
          <222> 493..493  
          <223> n = a, c, g or t  
  
25           <220>  
          <221> unsure  
          <222> 496..497  
          <223> n = a, c, g or t  
  
30           <220>  
          <221> unsure  
          <222> 500..500  
          <223> n = a, c, g or t  
  
35           <220>  
          <221> unsure  
          <222> 504..504  
          <223> n = a, c, g or t  
  
40           <220>  
          <221> unsure  
          <222> 506..507  
          <223> n = a, c, g or t  
  
45           <220>  
          <221> unsure  
          <222> 515..515  
          <223> n = a, c, g or t  
  
50           <220>  
          <221> unsure  
          <222> 520..520  
          <223> n = a, c, g or t  
  
55           <220>  
          <221> unsure  
          <222> 523..523  
          <223> n = a, c, g or t  
  
60           <220>  
          <221> unsure  
          <222> 525..526  
          <223> n = a, c, g or t  
  
65           <220>  
          <221> unsure  
          <222> 538..538  
          <223> n = a, c, g or t



<222> 441..441  
 <223> n = a, c, g or t

5       <220>  
       <221> unsure  
       <222> 448..448  
       <223> n = a, c, g or t

10       <220>  
       <221> unsure  
       <222> 457..457  
       <223> n = a, c, g or t

15       <220>  
       <221> unsure  
       <222> 525..525  
       <223> n = a, c, g or t

20       <220>  
       <221> unsure  
       <222> 539..539  
       <223> n = a, c, g or t

25       <220>  
       <221> unsure  
       <222> 564..564  
       <223> n = a, c, g or t

30       <400> 41  
 caacacaatc tcaaggtctt caaactggga tgcagaagag gcctctntca agttatgtng 60  
 tgctacttcc tgaacaaggc cttgaaaaga ntcaatggct tctngaagag cccgtaacca 120  
 cttcccgcta ttggctaggc ctttctcttt ctcttccaaa acagccttca tgggattcat 180  
 ctgtgcctct tttcctttta ataagttcyg aagctcctga actttggaag tgatagataa 240  
 cygttgagcc ttttcaagct ggacatttcc tatttcttct tttagagcct ttatttcccg 300  
 35   tttcaaatcc cgaacagttt tctccttggt agcaacattg agaagtgccg cctccagaag 360  
 ctcttctaca gacttgatct ttcacctctc tcatggatca ctctcttaag gtcttcaact 420  
 agagaggcaa aagaacccga ncatttgnta gttttangct cgaacttctt ttgtcagaga 480  
 tgaattttct gttcttattg cattcagctc ctcttcttta ggtgncacca gaataaggnc 540  
 agtgtcaaga aattcttcca cagnct 566

40       <210> 42  
       <211> 878  
       <212> DNA  
       <213> Homo sapiens

45       <220>  
       <221> unsure  
       <222> 595..595  
       <223> n = a, c, g or t

50       <220>  
       <221> unsure  
       <222> 756..756  
       <223> n = a, c, g or t

55       <220>  
       <221> unsure  
       <222> 766..766  
       <223> n = a, c, g or t

60       <220>  
       <221> unsure  
       <222> 815..815  
       <223> n = a, c, g or t

65       <220>  
       <221> unsure

<222> 818..818  
 <223> n = a, c, g or t

5      <220>  
       <221> unsure  
       <222> 827..827  
       <223> n = a, c, g or t

10     <220>  
       <221> unsure  
       <222> 831..831  
       <223> n = a, c, g or t

15     <220>  
       <221> unsure  
       <222> 853..853  
       <223> n = a, c, g or t

20     <220>  
       <221> unsure  
       <222> 866..866  
       <223> n = a, c, g or t

25     <400> 42  
 accttaacct cctctgacct gaagttccag ccatacatte ttttgccttt ttttcaatcc 60  
 atgcaaccat tcaccataac tcaaattaga agggacagac acctttggaa ataatttttt 120  
 gagaacttct ttagcctcca actcaacagc ttccacctgt tgctgccttt cattgttttt 180  
 ctctctctgg tgttcaactg catccttcaa agaattctaac tcattccaga gaccacttat 240  
 ttctttctct ctttctgaaa ttacttttaa taattcttca tgagggggaa aagaagatgc 300  
 30     ctggttggtg tttgttgtt taagctgctc aatttgggac ttaaacaatt tgttttcatc 360  
       ttgtacatcc tgtaacagct gtgttttgct agaaagatca ctctccctct cttttagcat 420  
       ggcttctaac ctcttcaatt cattttcctt ttctttcaac acaatctcaa gttcttcaaa 480  
       ctgtgatgca gaagaggcct ctttcaagtt atgttgtgct acttcttgaa catgtgcttt 540  
       taaaagattc attttcttct tgaagatcct gtaaccactt ccctgtattg gctangtctt 600  
 35     tctctttctc ttccaaaaca gccttcaagg tattcatctg ttctctttt ccttttaaaa 660  
       agttctgaag ctctgaact ttggaagtga tagataactg ttgagccctt tcaagcctgg 720  
       acattcctaa ttctctcttt aaaagcctta attccngttt caaaaanccga aacagttttc 780  
       ccccttgtaa gcaactttga gaaattcggc tccanaancc ctcttanaga nttgatcttt 840  
       caacctttcc aanggattac tttccnaaag ttccttaa 878

40     <210> 43  
       <211> 620  
       <212> DNA  
       <213> Homo sapiens

45     <220>  
       <221> unsure  
       <222> 391..391  
       <223> n = a, c, g or t

50     <220>  
       <221> unsure  
       <222> 400..400  
       <223> n = a, c, g or t

55     <220>  
       <221> unsure  
       <222> 411..411  
       <223> n = a, c, g or t

60     <220>  
       <221> unsure  
       <222> 414.. 414  
       <223> n = a, c, g or t

65     <220>  
       <221> unsure

<222> 418..418  
<223> n = a, c, g or t

5  
<220>  
<221> unsure  
<222> 421..421  
<223> n = a, c, g or t

10  
<220>  
<221> unsure  
<222> 427..428  
<223> n = a, c, g or t

15  
<220>  
<221> unsure  
<222> 435..435  
<223> n = a, c, g or t

20  
<220>  
<221> unsure  
<222> 446..446  
<223> n = a, c, g or t

25  
<220>  
<221> unsure  
<222> 454..455  
<223> n = a, c, g or t

30  
<220>  
<221> unsure  
<222> 463..463  
<223> n = a, c, g or t

35  
<220>  
<221> unsure  
<222> 465..465  
<223> n = a, c, g or t

40  
<220>  
<221> unsure  
<222> 471..471  
<223> n = a, c, g or t

45  
<220>  
<221> unsure  
<222> 480..480  
<223> n = a, c, g or t

50  
<220>  
<221> unsure  
<222> 484..484  
<223> n = a, c, g or t

55  
<220>  
<221> unsure  
<222> 490..490  
<223> n = a, c, g or t

60  
<220>  
<221> unsure  
<222> 493..493  
<223> n = a, c, g or t

65  
<220>  
<221> unsure  
<222> 500..500  
<223> n = a, c, g or t

5           <220>  
           <221> unsure  
           <222> 514..515  
           <223> n = a, c, g or t  
  
 10           <220>  
           <221> unsure  
           <222> 524..524  
           <223> n = a, c, g or t  
  
 15           <220>  
           <221> unsure  
           <222> 539..539  
           <223> n = a, c, g or t  
  
 20           <220>  
           <221> unsure  
           <222> 568..568  
           <223> n = a, c, g or t  
  
 25           <220>  
           <221> unsure  
           <222> 573..574  
           <223> n = a, c, g or t  
  
 30           <220>  
           <221> unsure  
           <222> 581..581  
           <223> n = a, c, g or t  
  
 35           <220>  
           <221> unsure  
           <222> 600..600  
           <223> n = a, c, g or t  
  
 40           <220>  
           <221> unsure  
           <222> 604..604  
           <223> n = a, c, g or t  
  
 45           <220>  
           <221> unsure  
           <222> 608..608  
           <223> n = a, c, g or t  
  
 50           <220>  
           <221> unsure  
           <222> 616..616  
           <223> n = a, c, g or t  
  
 55           <220>  
           <221> unsure  
           <222> 619..619  
           <223> n = a, c, g or t  
  
           <400> 43  
 60   accttaacct cctctgaccc tgaattccag ccatacatto ttttgccctt ttttcaaate   60  
      catgcaacca ttcaccataa ctcaaattag aagggacaga cacccttgga aataattttt   120  
      tgagaacttc tttagcctcc aactcaacag cttccacctg ttgctgcctt tcattgtttt   180  
      tcttcctctg gtgttcaact gcacccctca aagaatctaa ctcatccag agaccactta   240  
      tttctttctc tctttctgaa attactttta ataattcttc atgaggggga aaagaagatg   300  
      cctgttggtg gttttgttgt ttaagctgct caatttgga cttaaacaat ttgttttcat   360  
      cctgtacatc ctgtaacagc tgtgttttgc naagaaagtn actccccctc nccnttanca   420  
 65   nggggtnnca acctncttaa atacanttcc tttnncttta aananaatct naggttcctn   480  
      caanggggtan tgnaaaaaan gcccttttca agtnnggttg tgcnaattcc tgaacatgng   540  
      ctttaagatc attttctctg aagatccngt acnnttcccg nattggtagg cttcccttcn   600

ttcnaaanag cttcangng

620

5           <210> 44  
            <211> 623  
            <212> DNA  
            <213> Homo sapiens

10           <220>  
            <221> unsure  
            <222> 25..25  
            <223> n = a, c, g or t

15           <220>  
            <221> unsure  
            <222> 409..409  
            <223> n = a, c, g or t

20           <220>  
            <221> unsure  
            <222> 412..412  
            <223> n = a, c, g or t

25           <220>  
            <221> unsure  
            <222> 414..414  
            <223> n = a, c, g or t

30           <220>  
            <221> unsure  
            <222> 423..423  
            <223> n = a, c, g or t

35           <220>  
            <221> unsure  
            <222> 426..426  
            <223> n = a, c, g or t

40           <220>  
            <221> unsure  
            <222> 435..435  
            <223> n = a, c, g or t

45           <220>  
            <221> unsure  
            <222> 444..444  
            <223> n = a, c, g or t

50           <220>  
            <221> unsure  
            <222> 450..450  
            <223> n = a, c, g or t

55           <220>  
            <221> unsure  
            <222> 454..454  
            <223> n = a, c, g or t

60           <220>  
            <221> unsure  
            <222> 459..459  
            <223> n = a, c, g or t

65           <220>  
            <221> unsure  
            <222> 490..490  
            <223> n = a, c, g or t



5       <220>  
       <221> unsure  
       <222> 508..508  
       <223> n = a, c, g or t  
  
 10       <220>  
       <221> unsure  
       <222> 511..511  
       <223> n = a, c, g or t  
  
 15       <220>  
       <221> unsure  
       <222> 515..515  
       <223> n = a, c, g or t  
  
 20       <220>  
       <221> unsure  
       <222> 524..524  
       <223> n = a, c, g or t  
  
 25       <220>  
       <221> unsure  
       <222> 545..545  
       <223> n = a, c, g or t  
  
 30       <220>  
       <221> unsure  
       <222> 555..555  
       <223> n = a, c, g or t  
  
 35       <220>  
       <221> unsure  
       <222> 564..564  
       <223> n = a, c, g or t  
  
 40       <220>  
       <221> unsure  
       <222> 600..600  
       <223> n = a, c, g or t  
  
 45       <220>  
       <221> unsure  
       <222> 611..611  
       <223> n = a, c, g or t  
  
       <400> 44  
 50       accttaacct cctctgaccc tgaantccag ccatacatte ttttgccttt ttttcaaate 60  
       catgcaacca ttcaccataa ctcaaattag aaggacaga cacctttgga aataattttt 120  
       tgagaacttc tttagcctcc aactcaacag cttccacctg ttgctgcctt tccttggaag 180  
       tcttgttcac tttgtcctgc agcatttttt cagttgatgc caatgcttcc attgcttccc 240  
       agtttttctc ccgaagggtca ttgtttttct tcctctggtg ttcaactgca tccttcaaag 300  
       aatctaactc attccagaga ccacttattt ctttctctct ttctgaaatt acttttaata 360  
       attcttcatg aggggggaaaa gaagatgcct ggtgggaatt tgggtggttna ancnggccaa 420  
       ttnggnacct aaaanatttg gttncactn gtanaccng taaaaactgt gttttgctaa 480  
 55       aaagataacn cccctctct ttaagcangg nttcnaacct cttnattcat ttccttttct 540  
       tcaanaaatc tcagntctca aacngggatg caaaagaggc tcttcaagtt aagggtgggc 600  
       acttccggaa natggccttt aaa 623  
  
 60       <210> 45  
       <211> 625  
       <212> DNA  
       <213> Homo sapiens  
  
 65       <220>  
       <221> unsure  
       <222> 392..392  
       <223> n = a, c, g or t

5           <220>  
            <221> unsure  
            <222> 408..408  
            <223> n = a, c, g or t

10           <220>  
            <221> unsure  
            <222> 417..417  
            <223> n = a, c, g or t

15           <220>  
            <221> unsure  
            <222> 421..421  
            <223> n = a, c, g or t

20           <220>  
            <221> unsure  
            <222> 424..425  
            <223> n = a, c, g or t

25           <220>  
            <221> unsure  
            <222> 437..437  
            <223> n = a, c, g or t

30           <220>  
            <221> unsure  
            <222> 439..439  
            <223> n = a, c, g or t

35           <220>  
            <221> unsure  
            <222> 456..456  
            <223> n = a, c, g or t

40           <220>  
            <221> unsure  
            <222> 467..467  
            <223> n = a, c, g or t

45           <220>  
            <221> unsure  
            <222> 469..469  
            <223> n = a, c, g or t

50           <220>  
            <221> unsure  
            <222> 473..473  
            <223> n = a, c, g or t

55           <220>  
            <221> unsure  
            <222> 475..475  
            <223> n = a, c, g or t

60           <220>  
            <221> unsure  
            <222> 496..496  
            <223> n = a, c, g or t

65           <220>  
            <221> unsure  
            <222> 499..500  
            <223> n = a, c, g or t

            <220>

<221> unsure  
 <222> 516..516  
 <223> n = a, c, g or t

5

<220>  
 <221> unsure  
 <222> 524..524  
 <223> n = a, c, g or t

10

<220>  
 <221> unsure  
 <222> 535..535  
 <223> n = a, c, g or t

15

<220>  
 <221> unsure  
 <222> 538..538  
 <223> n = a, c, g or t

20

<220>  
 <221> unsure  
 <222> 560..560  
 <223> n = a, c, g or t

25

<220>  
 <221> unsure  
 <222> 582..582  
 <223> n = a, c, g or t

30

<220>  
 <221> unsure  
 <222> 592..592  
 <223> n = a, c, g or t

35

<220>  
 <221> unsure  
 <222> 594..594  
 <223> n = a, c, g or t

40

<220>  
 <221> unsure  
 <222> 604..604  
 <223> n = a, c, g or t

45

<400> 45  
 accttaacct cctctgaccc tgaagtccag ccatacattc ttttgccttt ttttcaaadc 60  
 catgcaacca ttcaccataa ctcaaattag aagggacaga cacctttgga aataattttt 120  
 tgagaacttc ttagcctcc aactcaacag ctccacctg ttgctgcctt tcattgtttt 180  
 tcttcctctg gtgttcaact gcatccttca aagaatctaa ctcatccag agaccactta 240  
 50 tttctttctc tctttctgaa attactttta ataattcttc atgaggggga aaagaagatg 300  
 cctgttggtg gttttgttgt ttaagctgct caatttgga cttaaacaat ttgttttcat 360  
 cttgtacatc ctggtaacag ctgtgttttg cntagaaaag attactcncc ctctctnttt 420  
 ncannngggt caaccntna aattacattt cccttntctt aaacaanant ctnangttct 480  
 tcaaactgtg atgcanaann ggctctttca agttangttg tgcnacttcc tgaanatntg 540  
 55 ctttaaagat catttcttcn tgaatccgt aaccacttcc cngtatgggc angncttccc 600  
 ttcncttcaa aaaagcctca aggta 625

<210> 46  
 <211> 4416  
 <212> DNA  
 <213> Homo sapiens

60

<400> 46  
 gccgcacagg gttttatagg atcacattga caaaagtacc atggagtttt atgagtcagc 60  
 atattttatt gttcttattc cttcaatagt tattacagta attttcctct tcttctggct 120  
 tttcatgaaa gaaacattat atgatgaagt tcttgcaaaa cagaaaagag aacaaaagct 180  
 tattcctacc aaaacagata aaaagaaagc agaaaagaaa aagaataaaa agaaaagaat 240

	ccagaatgga	aacctccatg	aatccgactc	tgagagtgtg	cctcgagact	ttaaattatc	300
	agatgctttg	gcagtagaag	atgatcaagt	tgacacctgt	ccattgaatg	tcgttgaaac	360
	ttcaagtagt	gttagggaaa	gaaaaaagaa	ggaaaaagaaa	caaaagcctg	tgcttgaaga	420
5	gcaggtcatc	aaagaaagtg	acgcatacaa	gattcctggc	aaaaaagtag	aacctgtccc	480
	agttactaaa	cagcccaccc	ctccctctga	agcagctgcc	tcgaagaaga	aaccagggca	540
	gaagaagtct	aaaaatggaa	gcgatgacca	ggataaaaaa	gtggaaaactc	tcatggtacc	600
	atcaaaaagg	caagaagcat	tgcccctcca	ccaagagact	aaacaagaaa	gtggatcagg	660
	gaagaagaaa	gcttcatcaa	agaaacaaaa	gacagaaaaa	gtcttcgtag	atgaaccctt	720
10	tattcatgca	actacttata	ttcctttgat	ggataatgct	gactcaagtc	ctgtggtaga	780
	taagagagag	gttattgatt	tgcttaaacc	tgaccaagta	gaagggatcc	agaaatctgg	840
	gactaaaaaa	ctgaagaccg	aaactgacaa	agaaaatgct	gaagtgaagt	ttaaagattt	900
	tcttctgtcc	ttgaagacta	tgatgttttc	tgcaagtagg	gctctttgtg	ttgtagactt	960
	gctaaaggag	aagtctggtg	taatacaaga	tgctttaaag	aagtcaagta	agggagaatt	1020
	gactacgctt	atacatcagc	ttcaagaaaa	ggacaagtta	ctcgtctgctg	tgaaggaaaga	1080
15	tgctgctgct	acaaaggatc	ggtgtaagca	gttaaccagg	gaaatgatga	cagagaaaga	1140
	aagaagcaat	gtggttataa	caaggtatga	agactgctga	ggaacattag	aaaaggaaca	1200
	taatgtattt	caaaacaaaa	tacatgtcag	ttatcaagag	actcaacaga	tgcatgtaa	1260
	gtttcagcaa	gttcgtgagc	agatggaggc	agagatagct	cacttgaagc	aggaaaatgg	1320
	tatactgaga	gatgcagtc	gcaacactac	aaatcaactg	gaaagcaagc	agtctgcaga	1380
20	actaaataaa	ctacgccagg	attatgctag	gttggtgaat	gagctgactg	agaaaacagg	1440
	aaagctacag	caagaggaag	ttcaaaaagaa	gaatgctgag	caagcagcta	ctcagttgaa	1500
	ggttcaacta	caagaagctg	agagaagggtg	ggaagaagtt	cagagctaca	tcaggaagag	1560
	aacagcggaa	catgaggcag	cacagcaaga	tttacagagt	aaatttgtgg	ccaaagaaaa	1620
	tgaagtagag	agtctgcata	gtaagcttac	agataccttg	gtatcaaaac	aacagttgga	1680
25	gcaaaagacta	atgcagttaa	tggaatcaga	gcagaaaaag	gtgaacaaag	aagagtctct	1740
	acaaatgcag	gttcaggata	ttttggagca	gaatgaggct	ttgaaagctc	aaattcagca	1800
	gttccattcc	cagatagcag	cccagacctc	cgcttcagtt	ctagcagaag	aattacataa	1860
	agtgattgca	gaaaaggata	agcagataaa	acagactgaa	gattctttag	caagtgaacg	1920
	tgatcgttta	acaagtaaa	aagaggaaact	taaggatata	cagaatatga	atttcttatt	1980
30	aaaagctgaa	gtgcagaaat	tacaggccct	ggcaaatgag	caggctgctg	ctgcacatga	2040
	attggagaag	atgcaacaaa	gtgtttatgt	taaagatgat	aaaataagat	tgctggaaga	2100
	gcaactacaa	catgaaattt	caaacaaaaa	ggaagaattt	aagattctaa	atgacaaaaa	2160
	caaagcatta	aaatcagaag	ttcagaagct	acagactctt	gtttctgaac	agcctaataa	2220
	ggatgttgtg	gaacaaatgg	aaaaatgcac	tcaagaaaaa	gatgagaagt	taaagactgt	2280
35	ggaagaatta	cttgaaactg	gacttattca	ggtggcaact	aaagaagagg	agctgaatgc	2340
	aataagaaca	gaaaattcat	ctctgacaaa	agaagtcaa	gacttaaaag	ctaaagcaaaa	2400
	tgatcaggtt	tcttttgcct	ctctagttag	agaacttaag	aaagtgatcc	atgagaaaga	2460
	tggaagagtc	aagtctgtag	aagagcttct	ggagcgagaa	cttctcaaag	ttgctaacaa	2520
	ggagaaaact	gttcaggatt	tgaacacagga	aataaaaggct	ctaaaaagaag	aaataggaaa	2580
40	tgtccagctt	gaaaaggctc	aacagttatc	tatcacttcc	aaagttcagg	agcttcagaa	2640
	cttattaaaa	ggaaaagagg	aacagatgaa	taccatgaag	gctgttttgg	aagagaaaga	2700
	gaaagacctt	gccaatacag	ggaagtgtt	acaggtactt	caagaagaaa	atgaatcttt	2760
	aaaagcacat	gttcaggaag	tagcacacaa	taacttgaaa	gaggcctctt	ctgcatacaca	2820
	gtttgaagaa	cttgagattg	tggtgaaaga	aaaggaaaat	gaattgaaga	ggttagaagc	2880
45	catgctaaaa	gagagggaga	gtgatcttcc	tagcaaaaaca	cagctgttac	aggatgtaca	2940
	agatgaaaaa	aaattgttta	agtcacaaat	tgagcagctt	aaacaacaaa	actaccaaca	3000
	ggcatcttct	tttccccctc	atgaagaatt	atttaaaagta	atttcagaaa	gagagaaaga	3060
	aataagtggt	ctctggaatg	agttagattc	tttgaaggat	gcagttgaac	accagaggaa	3120
	gaaaaacaat	gaaaggcagc	aacaggtgga	agctgttgag	ttggaggcta	aagaagttct	3180
50	caaaaaatta	tttccaaagg	tgtctgtccc	ttctaatttg	agttatggtg	aatggttgca	3240
	tggatttgaa	aaaaaggcaa	agaatgtat	ggctggaact	tcagggtcag	aggaggttaa	3300
	ggttctagag	cacaagttga	aagaagctga	tgaatgcac	acattgttac	agctagagtg	3360
	tgaaaaaatac	aaatccgtcc	ttgcagaaac	agaaggaaat	ttacagaagc	tacagagaag	3420
	tggttagcaa	gaagaaaata	aatggaaggt	taaggctgat	gaatcacaca	agactattaa	3480
55	acagatgcag	tcatcattta	catcttcaga	acaagagcta	gagcgattaa	gaagcgaaaa	3540
	taaggatatt	gaaaatctga	gaagagaacg	agaacatttg	gaaatggaac	tagaaaaggc	3600
	agagatggaa	cgatctacct	atgttacaga	agtcagagag	ttgaaggcac	agttaaatga	3660
	aacactcaca	aaacttagaa	ctgaacaaaa	tgaagacag	aaggtagctg	gtgatttgca	3720
	taaggctcaa	cagtcactgg	agcttatcca	gtcaaaaaata	gtaaaagctg	ctggagacac	3780
60	tactgttatt	gaaaatagtg	atgtttcccc	agaaacggag	tcttctgaga	aggagacaat	3840
	gtctgtaagt	ctaaatcaga	ctgttaacaca	gttacagcag	ttgcttcagg	cggtaaacca	3900
	acagctcaca	aaggagaaag	agcactacca	ggtgttagag	tgaagtaatt	gggaaactgt	3960
	tcatttgagg	ataaaaaagg	cattgtatta	tattttgcc	aattaaagcc	ttattttatgt	4020
	tttcaccctt	tctactttgt	cagaaacact	gaacagagtt	ttgtcttttc	taatccttgt	4080
65	tagactactg	ttttaaagaa	ggaaaaaaaa	gaagccaactc	tgtagacacc	ttcagagttt	4140
	agttttataa	taaaaactgt	ttgaataatt	agacctttac	attcctgaag	ataaacatgt	4200
	aatcttttat	cttattttgc	tcaataaaat	tggtcagaag	atcaaagtgg	taaagacaat	4260

gtaaaattta	acattttaat	actgatgttg	tacactgttt	tacttaacat	tttgggaagt	4320
aactgcctct	gacttcaact	caagaaaaca	cttttttggt	gctaatagtaa	tcgggttttg	4380
taatggcgtc	agcaaataaa	aggatgctta	ttattc			4416